Mergers During the First and Second Phase of Globalization:

Success, Insider Trading, and the Role of Regulation

Inaugural Thesis

To Obtain the Doctoral Degree

of the Faculty of Economics

Eberhard-Karls-Universität Tübingen

Germany

Submitted by

Gerhard Kling

Born in Munich

2004

Dean: Professor Dr. rer. pol. Renate Hecker

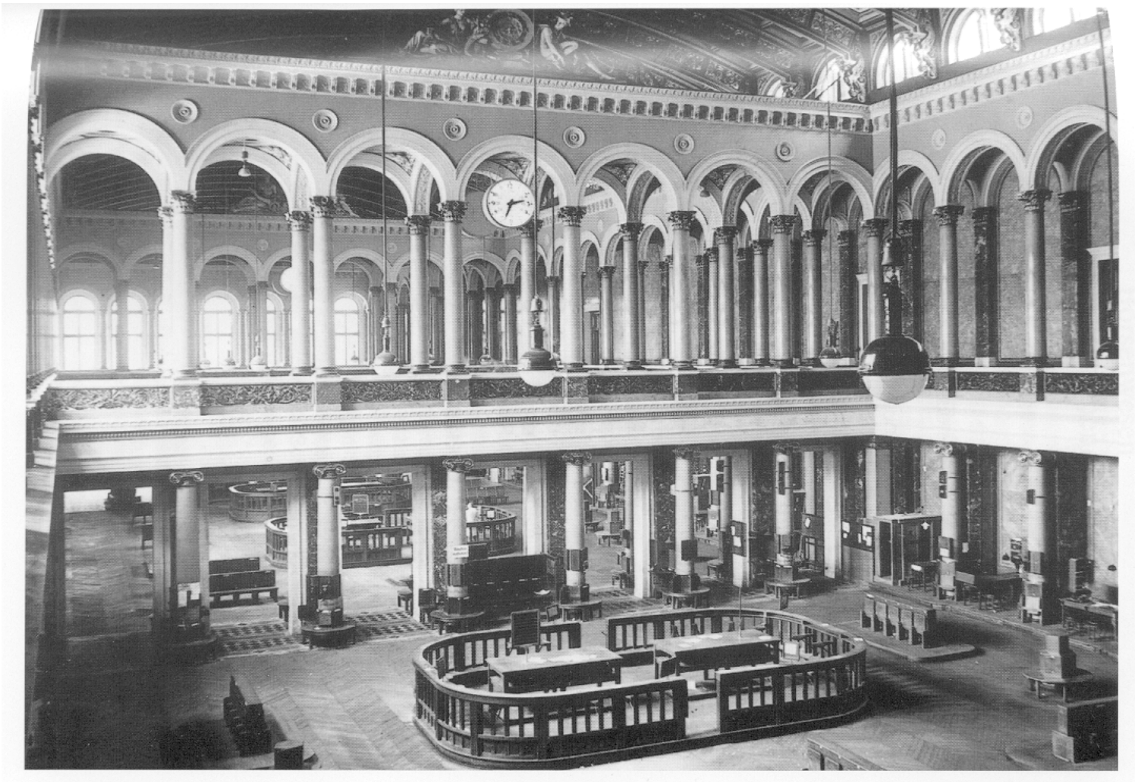
First Referee: Professor Dr. pub. oec. Jörg Baten

Second Referee: Professor Dr. rer. pol. Werner Neus

Day of the Disputation: 28th May 2004

Gerhard Kling

Mergers During the First and Second Phase of Globalization: Success, Insider Trading, and the Role of Regulation



"I find it difficult to think of economists and economic historians as separate animals. Their interests are fundamentally the same. The job of the economist is to explain how the economy works; the job of the economic historian is to explain how it worked in the past."

A. K. Cairncross

Acknowledgements

First of all, I take the opportunity to thank my advisor Jörg Baten for his steady support during `periods of suffering and doom´ that are typical for a dissertation project. He also encouraged me to attend conferences and present my research to an international audience. Even more important, he gave me the freedom to develop my own thoughts, make my own mistakes, and gain experience.

My thesis is based on five working papers – but they all contribute to one single story that highlights the success of mergers in Germany in different periods of time. Two of these working papers were presented at international conferences; hence, I want to thank for the many comments I received. My paper entitled `The impact of merger announcements on stock prices: The rejection of the merger paradox for German companies´ was presented at the annual meeting of the *Business and Economic Historical Society* in Memphis 2003. Lynne Pierson Doti, as my discussant and chair of the session, and Jari Eloranta gave me some very useful comments. This paper was also presented at the *5th European Historical Economics Society Conference* in Madrid 2003. Thus, I want to thank Stefano Battilossi, Albrecht Ritschl, and Joachim Voth for the challenging debates. In the general discussion, Stephen Broadberry asked whether mergers were successful in the long-run; now, I am able to respond, and my fifth chapter is dedicated to this task. My paper on the `Disclosure of mergers without regulatory restrictions: Comparing insider-trading in the year 1908 and 2000 in Germany´ was invited for presentation at the *NBER* conference on *Developing and Sustaining Financial Markets, 1820-2000*, which took place 2003 in Boston. I thank Marc Weidenmier for his excellent comments as well as Lance Davis, Larry Neal, and Eugene White who organized the conference and contributed remarkably to the high quality of discussions. I had also the chance to participate in the research seminar for PhD students organized by the *German Finance Association (DGF)* (2003 in Mainz) and in the *Third Summer School in Institutions, Economics, and History: Credit Networks and Economic Development* (2003 in Venice). On 26th January 2004, I presented the joint paper that I wrote together with Markus Baltzer entitled `Resiliency of the pre-World War I German stock exchange: Evidence from a panel vector autoregression´ at the *Workshop in Economic History* at the Humboldt University in Berlin. Furthermore, I took the chance to present my research several times at the *Economic Workshop* at the University of Tübingen. Consequently, I received many comments and suggestions to further improve my work. For their active role in the discussions, Werner Neus and Joachim Grammig deserve special thank.

The scholarship from the *German Research Foundation (DFG)* supported my research considerably and mitigated the financial burden imposed by attending international conferences. I also want to thank the *German Finance Association (DGF)*, the department of economic history of the Humboldt-Universität, and the *National Bureau of Economic Research (NBER)* for their financial support.

Furthermore, I thank the program committee that selected my paper on the \`Disclosure of mergers´ for the annual meeting of the *Economic History Society*. Due to the \`self-sacrificing´ work of my co-author Markus Baltzer our paper entitled \`Resiliency of the pre-World War I German stock exchange: Evidence from a panel vector autoregression´ attracts international interest; accordingly, we will present our paper at the *European Social Science History Conference* 2004 in Berlin. Even more noteworthy, this paper was also accepted for presentation at the *5th World Congress of Cliometrics* that will take place in Venice 2004.

The encouraging atmosphere in the graduate school and in our research group in economic history facilitated my work considerably. Henceforth, I thank all members of the graduate school and our research group for their friendship and steady support. I should emphasize, especially, the excellent comments due to Alexander Moradi and Aravinda Meera's advice with regard to language, spelling, style, and other \`horrible´ things.

Preface

Before you start reading, I want to highlight some basic convictions I have with regard to language, cited literature, and presenting problems inherent with data sets and empirical methods. I strongly believe in these convictions and I am convinced that they contribute to the quality of research.

The primary goal of scientific writing is to maintain clarity. Studying the books written by Day (1995, 1998) and Strunk and White (1999) helped to come closer to this goal – but it stays a difficult challenge. To avoid too wordy expressions, Day (1995, 1998) preferred to use active voice; thereby, `I´ and `We´ are interchangeably – but uncommon for German ears. By the way, I also learnt how to use semicolons in English – albeit I have still no idea how to use semicolons in German. Nevertheless, I hope that my language is simple enough to convey the sophisticated content.

Unfortunately, economic history is an extremely broad area of research that produces tons of literature every year; hence, I restricted myself to essential contributions published in refereed journals or outstanding edited volumes. Even worse, my research also includes topics in finance and econometrics which increases the related literature further. Accordingly, I cite only important sources that contribute to my research considerably. Generally, I focus on the `working paper style´ which avoids too lengthy reviews of literature. In contrast, my own data sets, methods, and results are of primary interest. Nevertheless, it is valuable to stress to what extent my research contributes to the existing strands of the literature. Of course, the typical literature in economic history usually contains hundreds of references; however the following quote that is due to William C. Roberts expresses my conviction best. “Manuscripts containing innumerable references are more likely a sign of insecurity than a mark of scholarship”.

Of course, empirical researchers often want to produce significant results, and only these results are usually published; however, having no results also contributes to science. Henceforth, I strongly believe that highlighting the limitations of data sets and methods is a crucial part of empirical research. For instance, I am very proud that I fail to detect a long-term impact of mergers on share prices and dividends. Thomas A. Edison put it in the following manner: “Results! Why, man, I have gotten a lot of results. I know several thousand things that won’t work.”

Often used symbols and abbreviations

P_t ....Stock price

E_{it} ....Error term

σ_e^2 ....Variance of the error term e which is equal to the residual of the CMR model

μ_i ....Mean return of stock I

R_{it} ....Observed daily returns

A ....Matrix that contains for all stocks i all observed daily returns for the whole estimation period

L ....Length of the estimation period

ε_{it}^* ....Abnormal return

τ_m or τ_n ....Indicate a specific point in time; thereby, $m \leq n$

$\hat{\mathbf{C}}(\tau_m;\tau_n)$ ....Vector of cumulated abnormal returns

$\bar{\varepsilon}_t^*$ ....Portfolio weighted abnormal return

$\overline{C}(\tau_m;\tau_n)$ ......Cumulated portfolio weighted abnormal return

D_{it} ....Dividends

N_{it} ....Nominal capital

$\mathbf{b}_{t-1}$ ....Estimated parameter vector for t; thereby, one uses only the information available at t-1

m_{it} ....Executed merger

$\Delta\mathbf{z}_{it}$..................Vector that contains the first difference in share prices and in dividends

inf_t ....Inflation rate at time t

Σ_j ....Coefficient matrix for lag j of the VAR in reduced form

$\mathbf{g}_t$..................Vector that contains unexpected shocks in inflation and growth rates

CMR ....Constant mean return model

MM ....Market model

CUSUM ....Cumulated sum of residuals

GARCH ....Generalized autoregressive conditional heteroscedasticity model

ARCH ....Autoregressive conditional heteroscedasticity model

ARIMA ....Autoregressive integrated moving average model

VAR ....Vector autoregression

Contents

1. Introduction

1.1 The aims of my dissertation project

1.1.1 Measuring the success of mergers

The merger wave that took place during the first phase of globalization, which lasted from 1895 to 1914,[1] changed the industrial structure in Europe and the U.S. remarkably. Therefore, it is of great importance to assess whether mergers were successful during this period. Noteworthy, studies that evaluate the success of mergers during the first phase of globalization are still lacking for Germany. One may argue that this statement is false and could refer to Huerkamp (1979). However, she defined success – by construction of her sample – as the ability of a firm to stay among the 100 largest companies.[2] Hence, shareholder value destroying mergers driven by `empire building´ that increase the firm size are seen as successful investments. Consequently, I totally disagree with her view. In contrast, I try to quantify the market response due to mergers and, hence, focus on the change in shareholder value. This imagination is in line with studies on the success of mergers for the United States and Great Britain. Generally, for the German case, economic historians concentrated on debates about the interrelation between the expansion of large scale enterprises, external growth, and mergers.[3] Maintaining size and survivorship were seen as major factors of success. But also `traditional´ cross-country studies showed that the large German enterprise was a main guarantee for superior economic development in the pre-World-War I period.[4] After reviewing new statistical material, however, the picture has to be corrected. A recent empirical cross-country study on that issue was written by Kinghorn and Nye (1996). They found evidence that German firms and production facilities were smaller compared to U.S. or French companies.[5] In addition, the concentration process was less developed in Germany. Besides these astonishing results, additional doubts emerge regarding the alleged success of large firms. In several empirical studies, Baten (2001 a, b, c) showed that small firms

---

[1] A subsequent session will discuss the so called `first phase of globalization´ in greater detail.

[2] Note that she used the data set collected by Kocka and Siegrist (1979); thereby, the 100 largest firms in 1907 were included – but their merger activity from 1887 to 1907 was studied. Hence, only companies that stayed among the 100 largest firms during this period were considered. Tilly (1982) argued that success can be defined in their study as the ability to stay among the 100 largest companies over the whole period.

[3] Tilly (1982, 1986), Gerschenkron (1962), Huerkamp (1979), and Feldenkirchen (1988) discussed the concentration process in different lines of business.

[4] I refer to the most famous contribution made by Chandler (1990). For late Victorian Britain, Elbaum and Lazonick (1986) argued in favor for large enterprises that had the ability to adopt the modern form of corporation; thereby, they referred to the German and U.S. example.

[5] Kinghorn and Nye (1996) calculated the number of workers in different industries for Germany, U.S., and France; thereby, German companies are relatively small. Especially, in the iron and steel industry, U.S. companies employed on average about two times more workers than German counterparts.

exhibited a larger total factor productivity. Moreover, he provided evidence that contradicted the usual opinion suggesting a steady increase in firm size between 1895 and 1912. In contrast, he found that the median of firm size stayed unchanged over time.[6]

When one turns to studies for the United States of America or Great Britain the scope regarding mergers is totally different compared to the `traditional´ research conducted in Germany. For instance, Leeth and Borg (1994, 2000) who covered the years 1905 to 1930 measured the economic impact of mergers by applying event-study methodology. In their study, successful mergers should yield an upsurge in market value.

Accordingly, my first aim is to assess the success of mergers based on the market response caused by merger announcements; thereby, a higher market value is the recipe for success and not firm size. Encouraged by the results of Baten (2001 a, b, c), I also collect data on mergers among smaller companies, which was, thus far, not done. Of course, my research contributes to close the data gap for Germany that is due to the absence of sources like Nelson (1957) and Eis (1971) who systematically collected data on mergers among U.S. companies.

1.1.2 Who gains from mergers?

If I, indeed, detected an increase in market values stemming from a merger announcement, another question would arise. Which type of shareholder gains from higher market values? Focusing on two types, namely insiders and outsiders, my aim is to answer this question; thereby, the so called run-ups prior to merger announcements serve as a measure for insider gains. Run-ups are changes in stock prices triggered by an impending merger announcement. As long as the merger is not yet public information, significant changes before the public release serve as a hint for insider-trading. If a market participant has only access to public sources like the official newspaper announcement, this participant belongs to the group of outsiders. In contrast, insiders possess private information; hence, they already know that a firm will announce publicly that they engage in merger activities. This superior knowledge leads to trading activities of insiders before the public announcement. Through this insider trading the private information is conveyed; thus, the market price is significantly influenced. Keown and Pinkerton (1981) used this measure to uncover insider activities around revealed mergers occurring in the years 1975-1978. Banerjee and Eckhard (2001) provided evidence for insider-trading in the year 1896-1903 known as the first merger wave. Both studies concentrate on the U.S. case.

---

[6] Due to data availability, his research is restricted to `Baden´; however, his empirical evidence can also be interpreted as a general statement regarding firm size. An exception, however, are some regions in which large scale enterprises (iron and steel, mining etc.) predominated the industrial structure.

Lacking regulatory restrictions are responsible for the appearance of two different forms of disclosure in the pre-World War I period in Germany. Some firms announce mergers after these mergers have already been executed, and others declare their desire to merge before the transfer of assets. Thus, one consideration is to assess whether the way of disclosure influences the gains respectively losses for insiders and outsiders. By comparing the pre-Word-War I period with mergers that took place in the year 2000 in Germany, I try to shed some light on the impact of regulations on insider activities and the ability of legislative restrictions to protect outsiders from insider trading.

1.1.3 Methodological issues and alternative approaches

Besides discussing economic issues like the change of shareholder value caused by mergers, I also try to thoroughly highlight methodological concerns. Consequently, I extend my former event-studies and conduct consistency checks to prove whether market frictions like non-synchronous trading affect my results. Caused by many restrictive assumptions inherent with event-study methods, I propose an alternative transfer function model. This time series approach has the capability to detect an \`empirical announcement day´; hence, this model serves as an alternative to identify run-ups. Furthermore, the microstructure of the Berlin stock exchange around 1900 can be explored by determining periods during which no trade is executed. These non-synchronous trading causes frictions that may influence my former results. Consequently, my thesis should also contribute to solve methodological issues.

1.1.4 The long-term impact of mergers

Using event-studies, I concentrate, thus far, on short-term market reactions caused by mergers. My additional concern is to shed some light on the long-term impact of mergers; thereby, an event-study approach must be replaced by more sophisticated methods. These superior models belong to the group of vector autoregressions (VAR). Besides focusing on mergers and, thus, micro-level shocks, I regard macroeconomic fluctuations as additional source of uncertainties. My panel VAR identifies the dynamics in share prices, dividends, and nominal capital caused by different kinds of shocks. In contrast to my short-term analyses, my long-run study covers the period from 1870 to 1913; thereby, I collected annual data. Changes in the regulatory environment at the beginning and in the middle of this period – especially the establishment of the new exchange law in 1896 – make the investigation promising from an institutional point of view.

1.2 Why should one care about historical evidence?

At conferences it is just a matter of time that the question arises whether we should care about mergers that occurred approximately 100 years ago. Or to put it differently, how can historical evidence improve our understanding regarding current mergers that typically take place during merger waves. I think the sense for looking at historical data is threefold. First, as mentioned, mergers normally occur in waves that affect single industries or the economy as a whole. The first merger wave can be located around 1900 and was a worldwide phenomenon. Collecting data on mergers during the first merger mania may help to clarify the emergence of the last merger wave that occurred shortly before the `bubble´ burst in the year 2000.

Second, globalization is not a new experience and makes a comparison between the first and second phase of globalization promising. Sachs and Warner (1996)[7] pointed out that already at the end of the nineteenth century a liberal international economic order allowed a global capitalism. This phase of liberalization is to some extent comparable to the current advances of a global economy. When one considers labor mobility and migration, the late nineteenth century was much more globalized than the late twentieth century.[8] The urgent question is whether the globalization about 100 years ago is similar to the integration process nowadays.[9] Besides other discrepancies, the emergence of multinational enterprises and the importance of intra-industry trade is a historically new experience as mentioned by Bordo et al. (1999). Caused by companies that sliced up their value chain by different sorts of foreign direct investment,[10] intra-industry trade becomes predominant in the trade between developed and developing countries. In contrast, in the late nineteenth century, the periphery mainly exported primary goods to developed countries. Assessing the degree of financial integration is also important for my research. However, the outcomes are ambiguous. Vásquez (2000) stressed[11] that the world is nowadays less financially integrated compared to the situation 100 years ago if one uses net capital flows as measure. In contrast, Bordo et al. (1999) and also the experiences from the Asian crises in 1997 showed that capital flows nowadays react much faster. This is due to the increasing importance of portfolio investments.

Consequently, the debate on the similarities and differences between both phases of globalization is still enduring; hence, my research will contribute some evidence on merger activities to this overall discussion.

---

[7] Williamson (1996) and Tilly (1999) besides others coined the term `first phase of globalization´; however, the exact location of this period exhibiting an increase in international integration is disputable.

[8] Hatton and Williamson (1994, 1998) discussed the mass migration from Europe.

[9] A current study on this subject by Bordo et al. (1999) discussed some issues and concluded that the current globalization of commodity and financial markets is historically unprecedented.

[10] This may also include cross-border mergers.

[11] This argument is based on figures obtained from the World Economic Outlook (1997).

Third, I can observe the `natural´ behavior of firms acting in an environment without regulatory restrictions. This historical experiment enables to quantify institutional changes and to derive policy recommendations whether state interventions are welfare creating – or not. Empirical investigations can also figure out why some firms behave nicely while others cheat. The third chapter, especially, deals with this advantage of analyzing historical periods.

1.3 Is this applied econometrics, finance or economic history?

Generally, the success of mergers and acquisitions is a central topic in corporate finance; however, I work with historical data, which points clearly into the direction of economic history. In addition, the econometric content of my thesis goes obviously far beyond `standard´ contributions in economic history. Therefore, someone may ask which label my research should have. Out of my point of view, my thesis combines different areas of specialization and may be regarded as a `jagged´ alliance among economic history, corporate finance, and applied econometrics. Nevertheless, my research can be characterized as quantitative economic history and cliometrics respectively. Hence, someone who expects only case studies and qualitative discussions might be disappointed. In contrast, my thesis is oriented toward empirical research and favors a more or less rigorous treatment of the applied econometric techniques.

1.4 The structure of my dissertation

My thesis can be split into two major parts; thereby, the second, third, and fourth chapter discuss the short-term economic impact of merger announcements. In the fifth chapter, I turn to the long-lasting influence of mergers on company characteristics like share prices, dividends, and the nominal capital. More precisely, the second chapter tackles the challenge to measure the market response triggered by merger announcement by relying on daily returns. Using daily returns is new for the pre-World-War I period. Thereafter, chapter three quantifies the scale of insider trading and the role of regulation. Comparing different regulatory frameworks established over time in Germany, I can assess the effectiveness of regulations. Observing unrestricted firm behavior allows to figure out whether a mechanism of self-regulation may work or the state should intervene. Although event-studies are very useful and widely applied to detect short-term market reactions, there may be alternative approaches. Hence, chapter four introduces modified transfer function models that overcomes usual pitfalls of event-studies. Besides the short-term market reactions, the long-term impact of mergers attracts my interest. It turns out that this task is highly challenging from an econometric point of view; thus, it deserves a thorough and rigorous analysis in chapter five. As a `waste product´ of the fifth chapter, several models are developed that clarify the expansion of enterprises and the decision to undertake a merger. A broad discussion of my results and an outline of future research topics conclude my thesis.

2. The impact of merger announcements on stock prices

2.1 Extended abstract

The scope of this chapter is to capture the market response triggered by merger announcements. Hence, the stock market decides whether a merger can be regarded as success. The construction of my short-term investigation enables to evaluate the performance of acquiring and target firms. Consequently, I can answer the question whether the merger paradox is observable in the pre-World-War I period in Germany. An event-study method applied to daily returns rejects the merger paradox based on my data. In addition, the adaptation process of stock prices according to newly available information is finished within a few days around the event day. Correspondingly, the exchange seems to be highly informationally efficient. To detect what affects the success of mergers, I use cumulated abnormal returns as dependent variable in a cross-sectional study. Controlling for direct and indirect effects, I construct a simultaneous equation approach. I uncover that banks exhibit remarkably positive abnormal returns in comparison to other lines of business.

2.2 Introduction

2.2.1 Former studies on the merger paradox

Thus far, a study that uses market valuations to quantify the success of mergers is still missing for the pre-World-War I period in Germany. Encouraged by the excellent availability of data sources[12] on executed mergers during this period in the United States, several event-studies for the U.S. were published. The analysis conducted by Leeth and Borg (1994, 2000) who covered the period from 1905 to 1930 attracted international interest. Furthermore, for the so called second merger wave that took place in the 1920s, studies did already exist.[13] Nevertheless, the majority of the literature in empirical finance started to measure the performance of acquiring and target firms around public releases of mergers occurring in the 1960s. Generally, evidence for the 1980s and later periods suggested that only the shareholders of target firms gain from takeovers and the share prices of acquirers are nearly unaffected.[14] Moreover, some empirical studies, for instance Travlos und Papaioannou (1991), found even negative abnormal returns of the acquiring firms. Why should firms

[12] As mentioned in chapter one, Eis (1971) and Nelson (1959) are by far the most cited sources for the United States.
[13] See Borg et al. (1989).
[14] Jarrell and Poulsen (1989) found positive abnormal returns for target firms – but negative or insignificant abnormal returns for acquiring firms.

initiate mergers if this means a loss for their shareholders? This finding is often called the `merger paradox´; however, there are several plausible solutions to clarify this puzzle.

One widespread idea is that institutions like monopoly commissions or more general the legal framework play a crucial role whether mergers create shareholder value. Jarrell and Poulsen (1989) pointed out that the lacking restrictions in the 1960s and 1970s compared to the 1980s are mainly responsible that the merger paradox can be confirmed for the latter period. However, compared to the pre-World-War I period, the 1960s or 1970s are highly `over- regulated´.[15]

The introduction of junk bonds in the 1980s facilitates the access to capital and, thus, is also seen as an essential factor to provoke a negative market reaction after merger announcements. If it is too simple to raise up money to finance an external expansion, inefficient mergers are more likely to be executed. One should take into consideration that larger companies have advantages in financing a merger. Correspondingly, Jarrell and Poulsen (1989) tried to control for this issue by considering the ratio of the firm size between the acquirer and the target firm. Besides the access to capital, they claimed that the larger the acquirer relative to the target the less important is the acquisition and, hence, the less likely is a negative market response.

The moral hazard problem inherent with the separation of ownership and control is regarded by many authors as additional source for the merger paradox. Thus, Jarrell and Poulsen (1989) argued that a manager favors even a shareholder value destroying merger to make the company larger. A larger company weakens the possibility for shareholders to control effectively the management. Shleifer and Vishny (1988) underlined that managers enjoy to increase their influence by `empire building´ that describes external growth for the sake to becoming larger – but not to increasing shareholder value. Based on these statements, there is also the imagination of the `market for corporate control´ that is due to Manne (1965). The prerequisite for this incentive mechanism is a highly positive correlation between the market valuation of a company and the quality of its management. If a management is inefficient, the risk of a tender offer[16] or a takeover bid will increase caused by the lower market value of the badly managed company. Accordingly, this external threat works as an incentive for the management to focus on the maximization of shareholder value.

[15] Borg et al. (1989) argued in a similar manner and favored their period 1919 to 1930 because it is less regulated. Nevertheless, my investigation period is even more liberal concerning mergers, cartels, and collusive behavior in general.

16 The SEC defines a tender offer as "(...) broad solicitation by a company or a third party to purchase a substantial percentage of a company's shares or units for a limited period of time."

How can my historical data set contribute to the understanding of the merger paradox? Compared to the 1960s and 1970s, the scale of regulation regarding horizontal mergers and collusive behavior was extremely low respectively did practically not exist in Germany during the pre-World-War I period. A monopoly commission, for instance, was not established, and other legal thresholds or local authorities were seldom an obstacle for mergers. More specifically, cartels and syndicates were part of the scene of the German industry,[17] albeit public opinion did not favor collusion.[18] In addition, political debates regarding cartels were quite common and intensified by events like the `coal need´ in 1900/1901 that led to a pronounced increase in coal prices.[19] Correspondingly, the first merger wave that occurred from 1898 to 1904 – based on the investigation by Banerjee and Eckhard (2001) – can be characterized as `mergers to monopoly´ as described by Stigler (1950).

Obviously, the separation of ownership and control was not predominant in the pre-World-War I period and, especially not, in small or medium sized companies. The manager of a smaller company was very often also the largest or at least an important shareholder. Furthermore, members of the advisory board had a considerable stake in the company; hence, typical free rider problems did not prevent effective control of the management. It is also noteworthy that the first concerns about incentive and control problems inherent with the separation of ownership and control were discussed by Berle and Means (1932) if one accepts the view of Scherer (1988). However, Pitelis (2004) pointed out that the incentive problems were already detected by Knight (1921) or are due to the "founding father of economics" Adam Smith (1776).

According to these highlighted discrepancies between the pre-World-War I period and the second phase of globalization, which starts after the second World War, one should expect that the merger paradox did not exist in the former period. This empirical finding would be also in line with former studies for the U.S. industry for the first merger wave[20] and the second merger wave in the 1920s. Thus, one may argue that providing evidence regarding the merger paradox for Germany is a contribution to fill an existing gap – but is from a methodological point of view nothing new. So what makes my analysis special?

[17] Fremdling and Krengel (1985) provided an excellent and critical overview on the importance of cartels in the German industry. They stated that the role of cartels is overstated regarding the impact on productivity, growth and price setting. They focused, however, mainly on the iron and steal industry.
[18] Gömmel (1985) argued that more than three fourth of the newspapers criticized the formation of cartels and syndicates. Especially, the leading newspapers measured by circulation like the `Morgenpost´ in Berlin attacked heavily the collusive behavior.
[19] Wengenroth (1985) described the general development of cartels in the German industry. A detailed discussion with regard to debates in parliament on collusive behavior was provided by Blaich (1973).
[20] Banerjee, Eckhard (2001), Leeth, Borg (1994, 2000), and Borg et al. (1989)

2.2.2 What makes my analysis special?

In contrast to all former event-studies for the pre-World-War I period, I apply an event-study method based on daily returns, which is in line with studies covering later periods. Borg et al. (1989) used monthly stock returns as well as Leeth and Borg (1994, 2000); however, Banerjee and Eckhard (2001) worked with weekly data – but this is still inferior compared to my precision in measuring market responses. Note that using monthly instead of daily returns makes it more difficult to detect abnormal stock price movements. Generally, the longer the chosen sampling interval the more cross-sectional units are needed to maintain a high power of the event-study. Morse (1984) presented a precise analysis on the usage of daily versus weekly respectively monthly returns.

Because I work on a daily frequency of my data, I have to determine the exact announcement day of a merger. This requires to read daily newspapers for a specific period of time – obviously, a time consuming task. Of course, daily newspapers are much more precise than weekly or monthly information sources that mainly focus on larger acquisitions. Correspondingly, when one wants to analyze the success of mergers among smaller companies, this is only possible by reading daily newspapers. Despite the fascinating stories on mergers spread by daily newspapers, using this historical sources comes with a cost. I have to restrict my analysis to a predefined time interval to be able to read the newspapers during this period. In a subsequent section, I will mention the pros and cons of this method of sampling.

Besides accumulating new information regarding smaller transactions, which indisputably is an interesting contribution to the strand of literature in economic history, I also develop a new econometric tool for event-studies. This tool, a simultaneous equation approach, controls for the influence of the choice of the estimation period on the inference of factors that lead to more successful mergers. Applying this new idea to other event-studies seems to be worthwhile.

2.2.3 The event-study method: A brief review

Since its introduction[21] into the field of empirical finance, the event-study method developed to one of the most often applied device to measure the economic impact of remarkable events. This many-sided method was used for a growing number of applications; moreover, Binder

[21] Especially Fama et al. (1969) and the study by Ball and Brown (1968) should be mentioned as pioneering applications of event-study methods.

(1985), Boehmer et al. (1991), Malatesta (1986), and Sefcik and Thompson (1986) developed some modifications of the basic event-study method.[22]

Economic theory offers explanations how economic events should influence the firms' market value. The aim of event-studies is to measure this impact by comparing the stock price movement in the presence of events with the normally expected price development. To achieve an accurate measurement, market prices should fully reflect currently available information. This strong informational efficiency of the market[23] is a crucial prerequisite; however, one can further relax this assumption by allowing an adaptation process of stock prices due to new information. This means that I have to define an event window during which the adaptation process should be finished.

2.2.4 The structure of this chapter

The remainder of this chapter is organized as follows. First, I thoroughly discuss the method of sampling; thereby, the pros and cons of my procedure are stressed. To give an impression regarding my data set, I present several descriptive figures and a selected number of brief case studies. Thereafter, I introduce the theoretical background of my event-study approach; thereby, I point out to what extent basic models should be modified to cope with special historical shortcomings. After making these modifications, I estimate the normal returns on the basis of the chosen estimation window by applying the constant-mean-return model (CMR).[24] Finally, I calculate the abnormal returns occurring in the event period and assess their significance. In a cross-sectional model, I try to figure out which explanatory variables influence the success of mergers and close with a brief discussion of my results.

2.3 The method of sampling

2.3.1 What can be regarded as a merger?

There are different kinds of transactions that could be called a merger, for instance subsidiary mergers or consolidations – but this investigation only deals with mergers after which the acquiring company survive in a legal manner, whereas the target firm becomes a part of the acquirer. The task to distinguish between a merger that fulfills this requirement and other forms of collusive behavior is sometimes tricky. During the pre-World-War I period and afterwards, several forms of non-tacit collusion existed,[25] Tilly (1982) stressed

---

[22] Armitage (1995) provided an excellent overview regarding the modifications and often used basic models.
[23] See Fama (1970), p. 383.
[24] See Masulis (1980).
[25] See also Feldenkirchen (1988).

the importance of pooling agreements (`Interessengemeinschaften´) that cannot be regarded as mergers. Nevertheless, such pooling agreements can be the starting point for a further integration of companies and, consequently, can lead to an actual merger. Besides being the potentially first step toward full unification, `pooled´ companies can together acquire a competitor. As an illustration, table 2.1 summarizes the newspaper articles that dealt with the merger of `Höchst´ and `Kalle & Co. AG´; thereby, the pooling agreement with `Leopold Casella & Co´ played a crucial role in financing the acquisition.

Table 2.1: The case of `Höchst´ and `Kalle & Co. AG´

<table>
<tr><th>Date of the newspaper announcements</th><th>Newspaper announcements in chronological order `Berliner Börsenzeitung´</th><th>Information provided by the year book `Handbuch der deutschen Aktiengesellschaften´</th></tr>
<tr><td>11th April 1908
Morning issue,
insert II</td><td>“`Höchster Farbwerke´ acquires `Kalle & Co. AG´ by issuing new shares. Already prior to this announcement rumors spread about an impending increase in nominal capital of `Höchster Farbwerke´”</td><td rowspan="7">“To deepen the relation between `Höchster Farbwerke´ and `Kalle & Co. AG´ an agreement was signed in 1908 to acquire shares from former principal shareholders. This acquisition was undertaken together with `Leopold Cassella & Co. GmbH´ and reached a volume of 4,000,000 Mark (nominal capital). The `Höchster Farbwerke´ now own shares with a nominal capital of 3,200,000 Mark, whereas `Leopold Cassella & Co. GmbH´ own 800,000 Mark in nominal capital”</td></tr>
<tr><td>12th April 1908
Sunday issue,
page 15</td><td>“`Höchster Farbwerke´ announces its annual accounts and also stresses that the management of the target firm should stay in charge after the acquisition of `Kalle & Co. AG´”</td></tr>
<tr><td>13th April 1908
Morning issue,
insert IV</td><td>“`Höchster Farbwerke´ convenes a shareholder gathering to decide about the issue of new shares and the acquisition. The gathering will take place on 9th May 1908”</td></tr>
<tr><td>19th April 1908
Sunday issue,
insert II</td><td>“The shareholder gathering of `Kalle & Co. AG´ will be held on 11th May 1908”</td></tr>
<tr><td>9th May 1908
Evening issue,
page four</td><td>“The shareholder gathering of `Höchster Farbwerke´ approves the increase of nominal capital by 10.5 million Mark. The nominal capital now reaches 36 million Mark.”</td></tr>
<tr><td>11th May 1908
Morning issue,
page 14</td><td>“The shareholder gathering of `Höchster Farbwerke´ accepts the acquisition of `Kalle & Co. AG´. A part of the acquired shares that represent 4 million Mark in nominal capital will be passed on to `Leopold Cassella & Co. GmbH´ with which a pooling agreement exists”</td></tr>
<tr><td>12th May 1908
Evening issue,
page 11</td><td>“Shareholder gathering of `Kalle & Co. AG´ approves the offer”</td></tr>
</table>

Source: The indicated issues of the `Berliner Börsenzeitung´ and the `Handbuch der deutschen Aktiengesellschaften´, issue for the years 1913-1914, volume I, page 1600-1602.

2.3.2 What can we learn from this case study?

I combine the information provided in the daily newspaper `Berliner Börsenzeitung´ and the year book `Handbuch der deutschen Aktiengesellschaften´[26] to get as much detail as possible about the initiated mergers. By reading the daily newspaper carefully, it is possible to capture the whole process of a merger that starts with the initial announcement and ends with the approval of the shareholders. In spite of the work intensity,[27] the `Berliner Börsenzeitung´ offers an excellent up to date information regarding events during the pre-World-War I period. For instance, the major results of the shareholder gathering of `Höchster Farbwerke´ that took place on 9th May 1908 were reported in the evening issue of the same day. Having in mind the long process and the many hurdles a merger has to overcome nowadays, the extremely short time span between the announcement, the call for the shareholder gatherings, and the approvals is astonishing. This high speed of decisions is a common feature of all of my detected mergers. Moreover, hostile takeovers or the replacement of the management were highly unusual in the pre-World-War I period. Typically, the newspaper announcements also contain information whether the management is allowed to stay in office or not. Generally, the quality and the high detail provided by the daily newspaper is remarkable. Accordingly, the daily information enables to precisely determine the announcement day, which is crucial for measuring the market response triggered by newly available information.

2.3.3 Drawing a sample

Caused by the time intensive and meticulous work inherent with reading daily newspapers, I had to restrict the time period. Accordingly, I included all mergers announced between 1st January 1908 and 31st June 1908 into my initial sample. Encouraged by Tilly's (1982) statements about drawing samples in historical time periods, this method is a widely accepted procedure.[28] The advantage is that all events are considered regardless if a firm is listed on the stock exchange respectively is very small. As a first step, I collect all relevant information and observe 101 announcements. However, to conduct an event-study, share prices have to be observed. This prerequisite together with the requirement that a sufficient amount of trades occurred reduce the number of included companies dramatically. Thus, I

[26] The `Handbook of the German Companies´ contains firm specific information on earnings, dividend payments as well as special activities, for instance, stock splits. Of course, the handbook does not contain any details about announcement days in daily newspapers.

[27] Note that the `Berliner Börsenzeitung´ had in these years a morning and an evening issue every working day; this makes the reading very time consuming. Even worse, the newspaper was also issued on Sundays.

[28] Tilly (1982) argued that choosing a specific time interval during which as much information is accumulated as possible is an appropriate method of sampling in economic history.

end up with forty-five cross sectional observations. Nevertheless, this sample size is in line with other event-studies, or simulation experiments executed by Brown and Warner (1980, 1985). Therefore, if one bases the analysis on daily returns, the econometric power of an event-study is very large.[29] Despite the high analytical quality, I should discuss the `historical power´ of my investigation, which is done in a subsequent section.

Moreover, it is crucial to determine precisely the day of the announcement to avoid false measurement; therefore, actuality and objectivity of the information source are important criterions. For the American market, the day of publication in the Wall Street Journal is usually used as event day. In Germany in the year 1908, the newspaper that satisfies these criterions best is the `Berliner Börsenzeitung´ because it possessed a great importance for investors and – thanks to telegram announcements – a high actuality. Therefore, the date of the first publication in the `Berliner Börsenzeitung´ is defined as the event day. Besides the determination of event days, more data are needed.

The `Berliner Börsenzeitung´ delivers stock prices of the Berlin stock exchange as well as other regional exchanges on a daily basis. This is an important improvement in comparison to using monthly or weekly stock returns,[30] which is too rough and leads to considerable methodical problems, for instance, cross correlation is more likely to occur.[31]

For the cross-sectional models additional information on stock characteristics like firm size is needed. A reliable source for company specific information is the `Handbuch der deutschen Aktiengesellschaften´. This year book also contains information on merger activities; however, only the year of an acquisition is usually reported.

After determining the event day, I turn to specify the event period that begins fifteen days before the announcement and ends fifteen days thereafter. In a later section, I justify this choice. Moreover, to estimate the normal returns, I collect fifty daily returns for each stock of my sample from the period January and February 1907. This estimation period is far enough away from the merger announcement and, hence, is not affected by the events, which is a prerequisite for estimating normal returns. Note that the estimated normal returns reflect the stock price movement without the merger event.

---

[29] See Morse (1984).

[30] See, for example, `Berliner Börsen-Courier´ and `Neuman's Kurs-Tabellen der Berliner Fonds-Börse´ that provide monthly data.

[31] Bernard (1987) found this result by running simulation experiments.

2.3.4 Why should I choose the year 1908 as investigation period?

Despite the high empirical power of my event-study, economic historians may wonder why I choose the year 1908 for my analysis. Henceforth, this decision needs a justification. Because only short time intervals can be analyzed, one should choose an interesting period for discussing the merger paradox. Note that my aim is to test whether the merger paradox exists and, accordingly, whether acquiring companies gain from mergers. Consequently, my null hypothesis states that the merger paradox can be observed. The rejection of the null hypothesis is more difficult in periods in which risky investments and, hence, mergers are punished by the market. If the market is bearish, such a punishment seems to be more likely.

Besides this argument, one can also point to the fact that in the year 1907, which serves as estimation period, higher normal returns should result.[32] Accordingly, in the downturn of the market in 1908, it should be more difficult to observe positive abnormal returns and, thus, to reject the \`merger paradox´. Figure 2.1 depict the \`Donnerindex´ to illustrate the basic trend in the market.[33] Although this index is often criticized for its composition, it should just give a first impression regarding the general market situation. In chapter five, I will construct my own market index and can overcome typical pitfalls of the \`Donnerindex´.

In addition, working with daily returns typically yields estimated normal returns that are not significantly different from zero. The following section provides my empirical findings for the normal returns and corresponding confidence intervals. In addition, chapter four deepens the discussion further and tests whether changes regarding the length or the location of my estimation period matter.

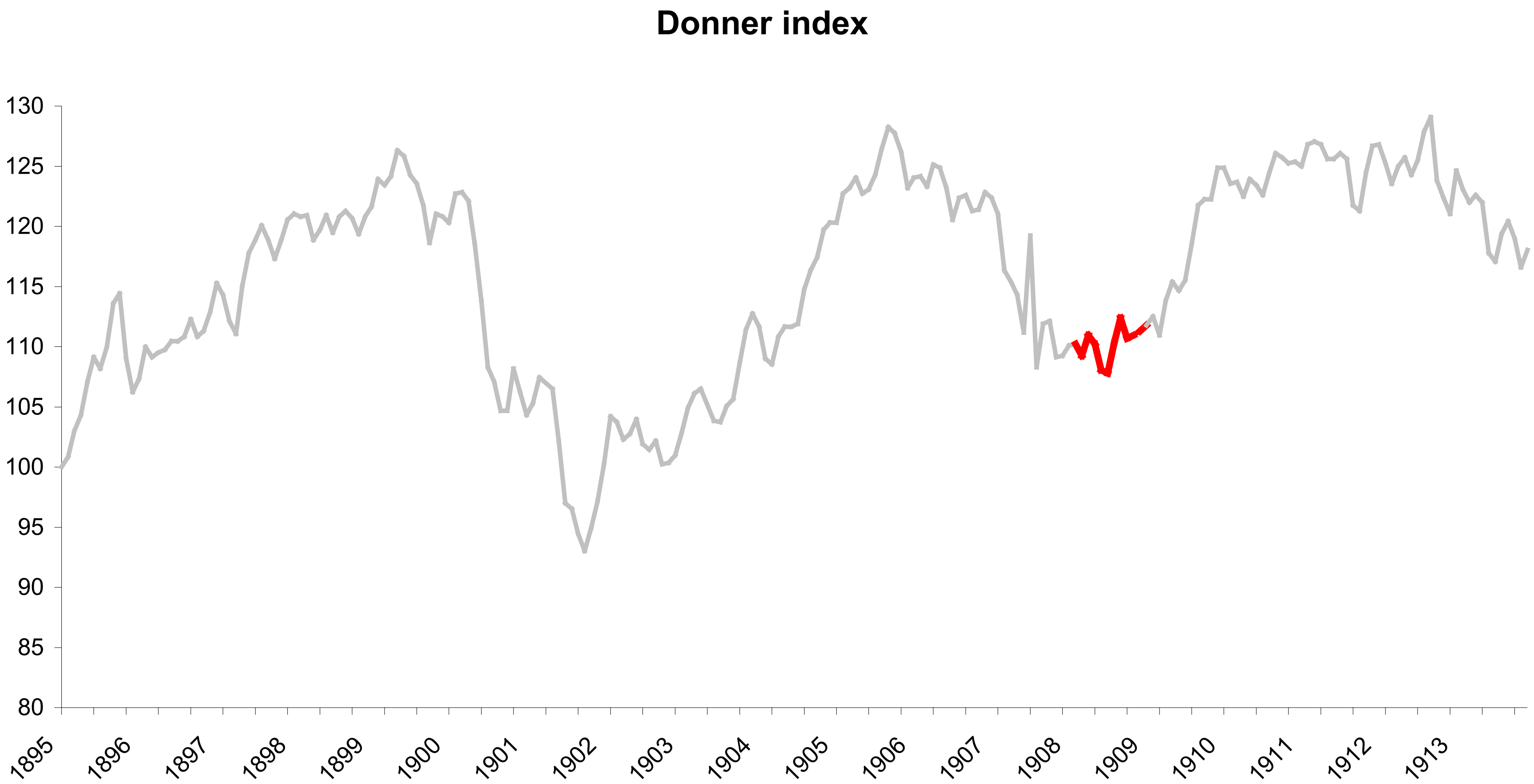
---

[32] Note that I used January and February 1907 as estimation period during which the \`Donnerindex´ reached values between 122.36 and 121.03. Thereafter, a pronounced decline set in, and the market lost more than 10% till the beginning of my event period; however, this decrease was not included in my estimation period. Furthermore, my fourth chapter provides evidence that switching the estimation period (using March and April 1907) or extending the period does not affect my results.

[33] For instance, Grabas (1992) listed the \`Donnerindex´ for this period in her data appendix; thus, I use this information to depict the market index in figure 2.1.

Figure 2.1: The `Donnerindex´ on a monthly basis 1895 to 1913

The event-period, 1st January to 30th June 1908, reaches a relatively low share price level compared to the previous estimation period, January to February 1907. On average, the estimation period exhibits 10% higher share prices in comparison to the event-period.



2.3.5 Descriptive analysis

To obtain a first impression of my data set, I divide the observed merger announcements into several subgroups according to line of business, success of merger, whether the management of the target firm changed after the takeover, how the deal is financed, the number of involved firms, and the motives behind the decision for initiating a merger, if this is published. Caused by missing values, I can only include 79 out of 101 cases in my descriptive analysis.

During my investigation period, I detect 101 mergers; however, the merger activity is far from being stable over the six considered months. Figure 2.2 shows that the peak is reached in April 1908. During this month 29.70% of all mergers are executed. Thus, I can conclude that the merger activity is time varying.

Figure 2.2: Time-varying merger activity over the six included month

Figure 2.2 depicts the number of mergers released during the respective week; thereby, the study starts in the first week of January 1908 and ends with the last week of June 1908.

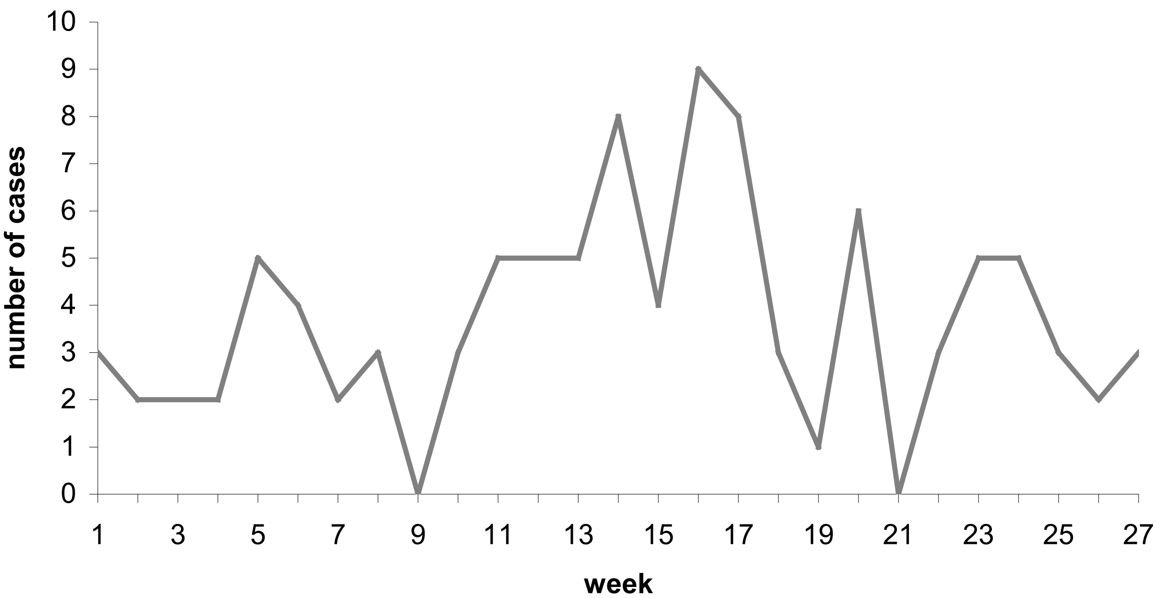


Table 2.2 uncovers discrepancies in merger activity that depend on the affiliation to a specific line of business. I make a crude distinction among the major groups in an economy, namely service, manufacturing, and raw material production. This is further refined; thereby, the subgroups are the largest groups in the respective sector. It is apparent that especially the banking industry is very active in undertaking mergers, whereas other industries like the mining sector exhibit only weak activities. This might be due to the formation of syndicates in the mining industry; accordingly, a subsequent section highlights this issue.

Table 2.2: Merger activity in different lines of business

Table 2.2 summarizes the number of mergers within each indicated line of business and the contribution, in per cent, of the respective category to the whole merger activities.

Line of Business	Absolute number of mergers	In per cent of total mergers
Banking	31	39.2%
Real estate	3	3.8%
Traffic	8	10.1%
Mergers in service industry	42	53.2%
Chemical	2	2.5%
Electrical engineering	3	3.8%
Brewery	4	5.1%
Other manufacturing industries	20	25.3%
Mergers in manufacturing	29	36.7%
Coal	3	3.8%
Potash	2	2.5%
Other raw material production	3	3.8%
Mergers in raw material production	8	10.1%
Total number of mergers	79	100%

In addition, I can determine the number of firms that are involved in a merger. In almost all cases only two firms interact; nevertheless, there are five out of 79 cases in which more than two firms merge. Due to the requirement that a company must be listed on a German stock exchange to observe daily returns, the number of observations drops to fifty. These companies are included in my event-study. However, lacking information on the details of the transaction not provided by the daily newspaper forces me to reduce the number of observations to forty-five for my cross-sectional models. Note that the way of financing a merger stays in five cases unfortunately unclear. Nevertheless, the merger paradox can be discussed by using the figures based on the event-study or by estimating the partial impact of being an acquirer in the cross-sectional analysis. Therefore, my analysis is by no means limited. Because the forty-five companies used in my cross-sectional model are of primary interest for my investigation, table 2.3 presents some descriptive figures. These figures include the line of business, whether the company is an acquirer or target, how the deal is financed, the decision of the shareholders and whether the management is replaced. Based on annual information provided by the `Handbuch der deutschen Aktiengesellschaften´, table 2.3 shows the market capitalization in million Mark. The thirteen target firms possessed an average market capitalization of 16.47 million Mark, whereas the 32 acquirers reached on average 46.45 million Mark. However, I cannot state that acquirers are about three times larger than their `victims´ in a transaction because not all targets are listed on the German stock exchanges.

Most announced mergers are successful; thereby, success means that the merger is executed after its public declaration. The advisory board and an extraordinary shareholder gathering must agree to the proposed merger. The minimum proportion of the shareholders that have to vote for the merger is defined by the statues of the specific company. Only two mergers fail to achieve the necessary majority.

It was uncommon to replace the management of the target firm after the merger, although the replacement of an inefficient management is often seen as one source of efficiency gains from mergers. Because only three cases can be observed in which the management is fired, the existence of a market for corporate control[34] and corresponding incentives for managers can hardly be supported by my data.

About 55.7% of the mergers were financed by cash payments, and the rest was conducted by using own shares as `acquisition currency´. This crude distinction seems to be a little bit misleading because cash payments are sometimes accompanied by an offer to transfer shares. Therefore, I regard a merger as being financed by cash payment if cash is the predominant payment – more than 90% of the total offer. Moreover, I observe that smaller acquisitions are more likely to be financed by cash payments, and if the target is not a listed company, cash payment is common.

2.3.6 The role of cartels and syndicates

I should stress that the formation of cartels and syndicates is not taken into account because they do not act like one firm after their formation. Despite this fact, cartels and syndicates play an important role for deciding to undergo a merger. Collusive behavior should reduce competition; therefore, one main motive to merge is lacking in industries in which cartels and syndicates prevail. Before suggesting that cartels reduce the driving force to merger, I should consider the spirit industry in which a syndicate called `Spirituszentrale´ existed. This syndicate possessed a strong market position especially in the southern regions of Germany. The `Nürnberger Presshefen- und Spiritusfabriken AG´ decided on 19th April 1908 to leave the syndicate. Because the management was afraid of retaliations, they bought two additional plants, one near Berlin, the other in Breslau, to increase firm size and to serve the local

[34] See Manne (1965).

markets in Northern Germany. According to this case study, the impact of cartels and syndicates cannot be determined in a clear manner.

Besides this `narrative´ evidence that the effect of cartels is unclear, econometric concerns make an evaluation of the partial impact less convincing. As far as I know, an unique measure for the degree of collusion that is applicable for every industry does not exist. Arguing in terms of average production cost and output prices achieved by a syndicate may work for the iron and steal industry[35] – but cannot be applied to the banking industry.

Moreover, also a practical concern arises because a considerable number of companies included in my cross-sectional models belong either to the banking or to the mining industry.

[35] Krengel (1982) provided several measures for the iron and steal industry regarding the degree of concentration and collusive behavior.

Table 2.3: Descriptive statistics for the 45 companies included into my event-study

Note that some acquirers like the `Osnabrücker Bank´ carried out several acquisitions; hence, they appear more than once in this table.

Name	Industry	Market capitalization in million Mark	Acquirer or target	Cash Payment or shares	Approval of shareholders	Replacement of management
Schleswiger Bank	Banking	44.87	TARGET	SHARES	Yes	No
Hattorf	Mining	3.48	TARGET	CASH	Yes	No
Osnabrücker Bank	Banking	20.24	ACQUIRER	CASH	Yes	No
Nürnberger Bank	Banking	7.26	TARGET	SHARES	Yes	No
Westlichen Boden-AG	Real estate	6.49	TARGET	SHARES	Yes	Yes
Handelsges. für Grundbesitz	Real estate	19.88	ACQUIRER	SHARES	Yes	Yes
Magdeburger Privatbank	Banking	39.89	ACQUIRER	SHARES	Yes	No
Wechslerbank	Banking	5.11	TARGET	SHARES	Yes	No
Bernburger Maschinenfabrik	Machinery	2.75	ACQUIRER	SHARES	Yes	No
Braunschweig. Maschinenfabrik	Machinery	0.36	TARGET	SHARES	Yes	No
Allgemeine Berliner Omnibusges.	Traffic	16.79	TARGET	SHARES	Yes	No
Osnabrücker Bank (2nd case)	Banking	20.30	ACQUIRER	CASH	Yes	No
Neptun	Shipyard	2.37	ACQUIRER	SHARES	No	No
Howaldtswerke	Shipyard	4.11	TARGET	SHARES	No	No
Schlesische Dampfergesellschaft	Traffic	2.53	TARGET	CASH	Yes	No
Essener Kreditanstalt	Banking	91.35	ACQUIRER	SHARES	Yes	No
Mülheimer Handelsbank	Banking	9.67	ACQUIRER	SHARES	Yes	No
Lindener Aktienbrauerei	Brewery	6.98	ACQUIRER	CASH	Yes	No
Deutsche Nationalbank	Banking	38.79	ACQUIRER	CASH	Yes	No
Bismarkshuette	Mining	28.53	ACQUIRER	CASH	Yes	No
Friedrichsegen	Mining	5.26	ACQUIRER	SHARES	Yes	No
Rheinisch-West. Disconto Ges.	Banking	105.49	ACQUIRER	SHARES	Yes	No
Telephonfabrik AG	Telephone	6.64	ACQUIRER	SHARES	Yes	No
Deutsche Steingutfabrik	Stone/Soil	0.49	ACQUIRER	CASH	Yes	No

Name Industry Market Acquirer or Cash Payment Approval of Replacement of

		capitalization in million Mark	Target	or shares	shareholders	management
Höchster Farbwerke	Chemical	166.89	ACQUIRER	SHARES	Yes	No
Danzinger Privat-Actien-Bank	Banking	9.56	ACQUIRER	CASH	Yes	Yes
Alexanderwerk	Mining	3.18	ACQUIRER	SHARES	Yes	No
Barmer Creditbank	Banking	1.80	TARGET	SHARES	Yes	No
Hildebrandsche Mühlenwerke AG	Grain mill	3.11	ACQUIRER	CASH	Yes	No
Bayerische Handelsbank	Banking	53.71	ACQUIRER	CASH	Yes	No
Wittener-Stahlroehren-Werke	Metal	6.56	ACQUIRER	SHARES	Yes	No
H. Renner & Co. AG	Chemical	7.78	ACQUIRER	CASH	Yes	No
Bayerische Handelsbank (2nd case)	Banking	54.55	ACQUIRER	CASH	Yes	No
Rheinische Stahlwerke	Metal	61.72	ACQUIRER	SHARES	Yes	No
Donnersmarkhuette	Mining	30.25	TARGET	CASH	Yes	No
Deutsche Nationalbank (2nd case)	Banking	38.78	ACQUIRER	SHARES	Yes	No
Eisenhuette Silesia AG	Mining	16.27	ACQUIRER	SHARES	Yes	No
Bayerische Handelsbank (3rd case)	Banking	54.20	ACQUIRER	CASH	Yes	No
Kasseler Federstahl	Metal	4.15	ACQUIRER	SHARES	Yes	No
Berliner Elektrizitätswerke	Electric	112.26	ACQUIRER	CASH	Yes	No
Elektrizitäts-Lieferungs-Ges.	Electric	25.23	TARGET	CASH	Yes	No
Terrainges. Berlin-Halensee	Real estate	65.78	TARGET	CASH	Yes	No
Dresdner Bank	Banking	246.71	ACQUIRER	CASH	Yes	No
Aschaffenburg Papierfabrik	Paper	2.39	ACQUIRER	CASH	Yes	No
Dresdner Bank (2nd case)	Banking	245.79	ACQUIRER	CASH	Yes	No

Henceforth, even if I was able to correctly quantify the scale of collusion, this hypothetical variable would only have about eight different realizations. This simply stems from the fact that cartels are an industry wide phenomenon and not necessarily company specific.

2.4 The theoretical background of event-studies

2.4.1 Random walk hypothesis

The basic idea is to consider stock prices P_l as following random walks. Note that the time is denoted by $l \in \{1,2,\ldots,L\}$ to indicate that I regard the estimation period and not the event period denoted $t \in \{1,2,\ldots,T\}$. A change in stock prices is only due to public information that can be seen as a white-noise process e_l. This process of newly available public information possesses the property that there is no autocorrelation.

Putting this feature in other words, it states that it is impossible to draw any conclusions from knowing the public information at l-1 that helps to improve the prediction for the public information occurring in the next period.

$$P_l = P_{l-1} + e_l \tag{2.1}$$

Thereby, the white-noise process e_l is characterized by a specific covariance function κ(j, l), and the mean function $\mu^e(l)$ is constant over time, but is allowed to differ from zero.

$$\kappa(j,l) = \begin{cases} 0 & \forall j \neq l \wedge j,l \in \{1,2,...,L\} \\ \sigma_e^2 & \forall j = t \wedge j,l \in \{1,2,...,L\} \end{cases} \tag{2.2}$$

Note that the variance σ_e^2 remains unchanged over time. Thus, it can be easily seen that taking the first difference from expression (2.1) yields a stationary white-noise process.

2.4.2 Merger announcements – events with great influence

Although the random walk hypothesis was criticized because it is partially refutable,[36] I maintain this hypothesis. It is appropriate to describe short-term price fluctuations taking place in a normal environment. These normal situations cover periods in which informed trading is relatively rare because meaningful events, which could have an essential impact on the true value of the underlying stock, do not impend. The public declaration of a merger obviously possesses an enormous impact on the true values of the interacting firms. Jennings (1994) showed that the information asymmetry increases around merger announcements. This is the breeding-ground for informed trading because knowing that a merger is going to occur

[36] See, for instance, Fama and French (1988). They uncovered negative autocorrelation of returns in the long-run; thus, the error term e_l in (2.1) would exhibit autocorrelation.

provides extraordinary profits that exceeds the costs of trading, especially transaction costs. One can identify two sources of informed trading.

Some insiders, for instance managers of the interacting companies, know with great confidence that a merger is going to be declared. Using their informational advantage, they start to buy stocks, if they expect that the true value exceeds the actual stock price. This insider trading reveals private information through the price process such that the market value narrows to the true value expected by the insiders. This adaptation of the market price to the true value, which is changed by the merger, takes maybe some days. The more market participants belong to the group of insiders the tougher is the competition among these insiders and the faster the adaptation process is finished.[37]

The other source of informed trading stems from above-average analytical skills that enables to achieve informational advantage from public information. This means that some market participants make more precise predictions about the possibility of a merger than others do.[38]

I highlight that the event study method can and should be applied in such situations in which the strong market efficiency is relaxed by allowing informed trading. Consequently, the market price P_l does not perfectly reflect all public and private information that exist at time l. However, the competition among insiders and their trading patterns yield to an adaptation process that guarantees that the private information is reflected in the market price P_l with a time lag. I tackle this problem by constructing a thirty days window surrounding the event day.

Moreover, the presence of informed trading and the described adaptation process are in conflict with my random walk hypothesis. To see this point, consider the second source of informed trading. As it is captured in the variance structure $\kappa(j, l)$, the public information e_{l-1} does not give a clue about e_l. However, the variance structure forbids that market participants have the capability to turn e_{l-1} in private information, for instance, they update after observing e_{l-1} their expectation about the probability that a merger occurs. Thus, if informed trading exists that is triggered by an impending considerable event like a merger announcement, the random walk hypothesis failed. This failure is exactly what I try to show in my event-study.

Accordingly, I use the random walk hypothesis as null hypothesis that the event is not meaningful; thus, the merger does not change the fundamental value of the interacting firms. Based on the random walk, the next section introduces a model that enables to determine the

[37] See Kyle (1985) for a theoretical model that describes the strategic behavior of insiders and the competition between them was modeled by Holden and Subrahmanyam (1992).
[38] See Kim and Verrecchia (1994) that provide a theoretical model of this sort of informed trading.

normal return. Note that the normal return determines the return that I would have expected if the merger announcement had not occurred. The economic impact of a merger is, thus, the return in the presence of the public declaration minus the normal return. Nevertheless, informed trading has other consequences discussed in chapter three.

2.4.3 The constant-mean-return (CMR) model

Masulis (1980) developed the CMR model that represents the basis of my model. Note that the CMR is nowadays not widely applied because the market model works better under normal circumstances. But finding an appropriate market index on a daily basis for German companies, is difficult in the year 1908, and even available monthly indices[39] are not generally accepted.

Taking the first difference of equation (2.1) and dividing by P_{l-1} provides the return of stock prices R_l.[40] My sample consists of n different stocks i and the estimation window is denoted as $l \in \{1,2,\ldots,L\}$; hence, I get the following expression.

$$R_{il} = \mu_i + e_{il} \quad \text{where: } R_{il} \equiv \frac{P_{il} - P_{il-1}}{P_{il}} \text{ and } E(e_{il}) = 0 \tag{2.3}$$

The stock return, like the first difference of equation (2.1), follows a white-noise process; thereby, e_{il} denotes a white-noise process with mean function that is equal to zero. Thus, μ_i represents the mean function of R_{il} which is supposed to be constant over time. For convenience, I put expression (2.3) in matrix notation; thereby, bold letters indicate matrixes.

$$\mathbf{R_l} = \boldsymbol{\mu} + \mathbf{e_l} \tag{2.4}$$

In equation (2.4) all vectors are column vectors with dimension n×1. Note that I maintain the random walk hypothesis and, hence, my null hypothesis that the event has no impact on share prices. Following this logic, I characterize in equation (2.4) the normal return. This is the return I would expect, if the event did not occur. This expression is the core of the constant-mean-return (CMR)[41] model. Estimating (2.4) is straightforward.

$$\hat{\boldsymbol{\mu}} = L^{-1} \cdot \sum_{l=1}^{L} \mathbf{R_l} \tag{2.5}$$

---

[39] A already mentioned the `Donnerindex´ - but also the more valid Eube (1998) index is currently not accepted, in general.

[40] This model can be expressed in natural logarithms correspondingly. In this case, the first difference is obviously an approximate return. However, the results of my event study are also valid if I use a log linear specification.

[41] Although the CMR is the simplest model Brown and Warner (1980,1985) found that it yields very similar result in comparison to more sophisticated models. The variance of the abnormal return is not significantly reduced using a more complex model such as multi-factor models. In chapter four, I provide additional evidence.

Consider L stands for the length of the estimation period, which is the same for all stocks. And $\mathbf{R_l}$ is a n×1 dimensional vector that collects for time point $l \in \{1,2,\ldots,L\}$ of the estimation window the return of each stock i. Sometimes it is more convenient to use the matrix form (2.6). I will switch between these two notations in the mathematical appendix.

$$\hat{\boldsymbol{\mu}} = L^{-1} \cdot \sum_{l=1}^{L} \mathbf{R}_l = L^{-1} \cdot \mathbf{A1} \tag{2.6}$$

I define matrix **A** as n×L dimensional matrix that contains all stocks i and for each time l all observed daily returns. I also define the unity vector **1** as being L×1 dimensional. Moreover, I derive the variance of the method of moments estimator directly from expression (2.5).

$$Var\hat{\boldsymbol{\mu}} = Var\left[L^{-1} \cdot \sum_{l=1}^{L} \mathbf{R_l}\right] = L^{-2} \cdot \sum_{l=1}^{L} Var\mathbf{R_l} = L^{-1} \cdot \boldsymbol{\sigma}^2 \tag{2.7}$$

At that point, note that successive returns of stock i are supposed to be uncorrelated over time. Therefore, it is possible to draw the variance operator under the sum operator. The resulting variance vector $\boldsymbol{\sigma}^2$ is obviously a n×1 dimensional vector that allows for differences in variances among stocks i and states that the variances remain unchanged over time $l \in \{1;2;\ldots;L\}$.[42] In comparison to the CMR model proposed by Masulis (1980), I use this variance expression to control for the inaccuracy of estimating normal returns.

2.4.4 Abnormal returns and their statistical properties

Having determined the normal return by the CMR, I can now define the abnormal return $\boldsymbol{\varepsilon}_t^*$ that stems from the merger announcement. The abnormal return is simply the difference between the observed return vector $\mathbf{R_t}$ during the event window $t \in \{1,2,\ldots,T\}$ and the part of this observed return that can be predicted using the CMR model.

$$\mathbf{R_t} = E(\mathbf{R_t}|\mathbf{A}) + \boldsymbol{\varepsilon}_\mathbf{t}^* \tag{2.8}$$

Note that **A** is the data matrix containing the daily returns of the estimation period. This data matrix is the ingredient for estimating the mean vector $\boldsymbol{\mu}$ like mentioned in (2.6). The conditional mean in equation (2.8) is obviously equal to $\hat{\boldsymbol{\mu}}$ which follows from taking the conditional expectation of equation (2.4) and replacing the mean vector by its estimate as described in expression (2.6). Therefore, I obtain an estimate for the abnormal return vector.

$$\hat{\boldsymbol{\varepsilon}}_\mathbf{t}^* = \mathbf{R_t} - \hat{\boldsymbol{\mu}} \tag{2.9}$$

[42] These assumptions are very convenient and show that my model share common features with share time series models provided by Dyckman et al. (1984).

Under the null hypothesis that the event has no economic impact, one can now derive the statistical properties of the abnormal returns.

Result 1[43]

Under the null hypothesis, the conditional distribution of the estimated abnormal return vector $\boldsymbol{\varepsilon}_{\mathbf{t}}^{*}$ after having observed the data matrix **A** and assuming that the abnormal returns are jointly normally distributed[44] can be described as follows.

$$\hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^{*} \sim N\left(\mathbf{0}; \mathbf{I}_{n\times n}\boldsymbol{\sigma}_{e}^{2} + \mathbf{I}_{n\times n}Var(\hat{\boldsymbol{\mu}})\right) \tag{2.10}$$

What I want to show is that the abnormal returns deviate significantly from this conditional distribution; correspondingly, the merger has an essential impact on the market value of the affected firms.

2.4.5 Aggregation of abnormal returns over time

Because the adaptation process of market prices caused by the merger announcement takes several days, it seems to be worthwhile using the aggregated value of the abnormal returns as a measure of the impact of mergers. The estimated cumulated abnormal return vector $\hat{\mathbf{C}}(\tau_m;\tau_n)$ with dimension n×1 covering the time period from τ_m to τ_n is defined in the following manner.

$$\hat{\mathbf{C}}(\tau_m;\tau_n) \equiv \sum_{t=\tau_m}^{\tau_n} \hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^{*} \tag{2.11}$$

Result 2 offers an appropriate test statistic for the cumulated abnormal return, which is discussed in detail in the mathematical appendix.[45]

---

[43] This is derived and explained in detail in the mathematical appendix. I deviate from conventional CMR models by taking the variance of the estimated mean vector into account. Thus, I control for an imprecise estimation.

[44] It is usually possible to relax this assumption and replace it by a parameter free representation – but the results are almost the same (see Corrado and Zivney, 1992). Moreover, this assumption simplifies the analysis considerably.

[45] The mathematical appendix highlights also some interesting insights with respect to an optimal choice of the event window.

Result 2

Assuming that the abnormal returns are jointly normally distributed, like in result one, yields the following test statistic that is t-distributed with T_1-2 degrees of freedom. T_1 represents the number of the days belonging to the selected event window from τ_m to τ_n. The estimated standard deviation appears in the denominator.

$$\frac{\hat{C}_i(\tau_m;\tau_n)}{\hat{\sigma}_i(\tau_m;\tau_n)} \sim t_{T_1-2} \qquad \forall i = \{1;2;...;n\} \tag{2.12}$$

Thereby, $\hat{C}_i(\tau_m;\tau_n)$ denotes the ith element of the $\hat{\mathbf{C}}(\tau_m;\tau_n)$ vector and $\hat{\sigma}_i(\tau_m;\tau_n)$ is the iith. element of the covariance matrix $\mathbf{V_c}$, which is derived in the mathematical appendix.

2.4.6 Cumulating abnormal returns over time and over cross sectional units

To assess whether the equally weighted portfolio of the firms, included in the sample, shows systematically higher returns from τ_m to τ_n, I use a modified version of cumulated returns. Accordingly, I now have to aggregate abnormal returns over time and cross sectional data. The first step is to estimate the sample average of the abnormal returns at time t and to determine the variance of this estimate. For this purpose result 3 provides the results that are discussed in the mathematical appendix in an accessible manner.

Result 3

The estimate for the sample average takes the following form; thereby, **1** is a n×1 dimensional unity vector.

$$\bar{\varepsilon}_t^{\,*} = n^{-1} \cdot \left(\hat{\boldsymbol{\varepsilon}}_\mathbf{t}^*\right)' \mathbf{1} \tag{2.13}$$

This expression is obviously the arithmetic mean of the abnormal returns at time t. If I assume that the abnormal returns at t are uncorrelated among securities i, I can write the estimated covariance matrix of $\bar{\varepsilon}_t^{\,*}$ in the following fashion.

$$Var\left(\bar{\varepsilon}_t^{\,*}\right) = n^{-2} \cdot tr\left(\mathbf{I}_{n\times n}\boldsymbol{\sigma}_\mathbf{e}^\mathbf{2} + \mathbf{I}_{n\times n} Var(\hat{\boldsymbol{\mu}})\right) \tag{2.14}$$

The second step carries out the aggregation over time. I denote the equally weighted cumulated abnormal return covering the time span from τ_m to τ_n with the term $\overline{C}(\tau_m;\tau_n)$. To test whether this cumulated abnormal return of the whole portfolio is significantly different from zero, I use the following test statistic.

Result 4[46]

The standardized $\overline{C}(\tau_m;\tau_n)$ is approximately standard normally distributed.

$$\frac{\overline{C}(\tau_m;\tau_n)}{\overline{\sigma}(\tau_m;\tau_n)} = \frac{\sum_{t=\tau_m}^{\tau_n} \overline{\varepsilon}_t^*}{\sqrt{n^{-2} \cdot T_1 \cdot tr\left(\mathbf{I}_{n\times n}\boldsymbol{\sigma}_{\mathbf{e}}^{\mathbf{2}} + \mathbf{I}_{n\times n} Var(\hat{\boldsymbol{\mu}})\right)}} \overset{a}{\sim} N(0;1) \tag{2.15}$$

With these results, I try to uncover the abnormal returns and cumulated abnormal returns that stem from the merger announcement and to assess their significance.

2.5 Empirical results of the event-study

2.5.1 Results for the estimation period

The CMR model yields the mean vector μ and the variance of these estimates; thereby, fifty observed daily returns during an estimation period are taken into account. To avoid biased normal returns that could stem from the impact of the merger announcements, I choose for my estimation the period from January to February 1907. This should be far enough away from the first announcement occurring in January 1908. To illustrate this procedure, figure 2.3 plots the upper and lower bounds of the estimated mean vector μ on a 95% level of confidence.

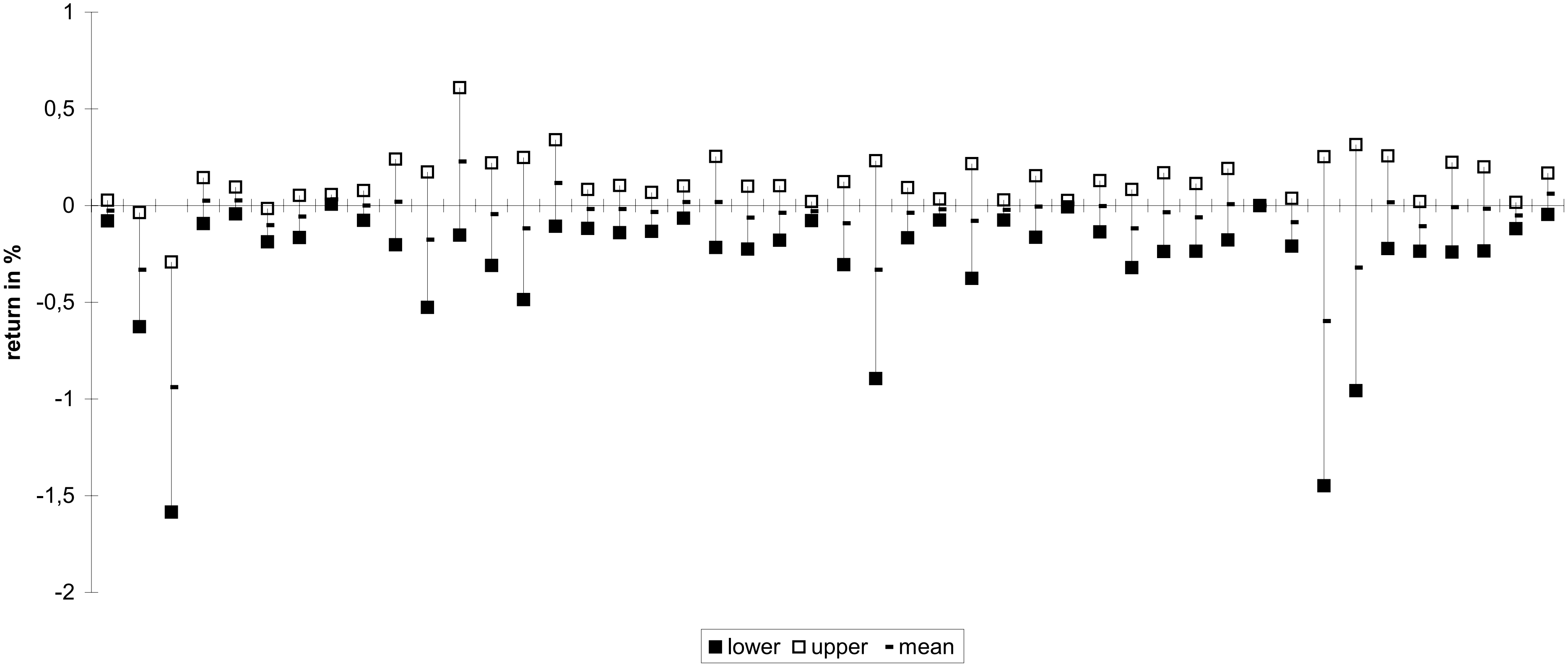
The null hypotheses that the mean return is equal to zero can rarely be rejected. Nevertheless, an exception is the mining company `Hattorf´, the third case in the diagram, that shows a negative drift during the estimation window.

Afterwards, I turn to the estimation of the variance vector $\boldsymbol{\sigma}_{\mathbf{e}}^{\mathbf{2}}$; thereby, I need the residuals from equation (2.4). Recall that the vector $\mathbf{R}_{\mathbf{l}}$ is the collection of the daily returns of all stocks at time l; thereby, l belongs to the estimation period $l \in \{1,2,\ldots,L\}$. Now, I can determine the distribution of the estimated abnormal return vector $\boldsymbol{\varepsilon}_{\mathbf{t}}^{*}$, using result 1. From this result, it is straightforward to obtain a test statistic, which enables to assess whether the estimated abnormal returns of stock i at time t are significantly different from zero.

[46] The mathematical appendix provides the details.

Figure 2.3: Confidence intervals of the estimated mean return for the whole sample of fifty stocks

Figure 2.3 plots the upper and lower bounds of a 95% confidence interval of the estimated mean returns using the period January to February 1907 as estimation window. Note that I depict the confidence interval for every stock in the sample; thereby, the order follows the chronology of announcements.



2.5.2 The justification for the choice of the event period

The design of the event window is one of the crucial problems of event-studies – but seldom discussed. If I choose an event period that possess a too large length T, it will be hard to detect significant cumulated abnormal returns defined over this period. Using result 4 and the information provided by the mathematical appendix, it appears that the larger T the larger $\overline{\sigma}(1;T)$, thus the more difficult to reject the null hypothesis $\overline{C}(1;T)$=0. In contrast, an event period that covers too few observations can be misleading because, as mentioned above, the adaptation process may not be finished yet. This corresponds with an inadequate measurement of the economic impact of an event. Therefore, a properly chosen event period is essential for obtaining reliable results; however, there exists no generally accepted method to determine what can be regarded as optimal event period. To give a hint whether my applied length T of the event period is useful, figure 2.4 plots the number of cases that exhibit significant abnormal returns at time t. The adaptation process seems to take place especially during the period ranging from eight days before to seven days after the public announcement. Leaving the core of my event period, the number of significant abnormal returns declines rapidly.

Because only the significant abnormal returns are counted in figure 2.4, I cannot make any statement about the significance level of the other neglected abnormal returns. The discrete structure and the use of the two exclusive conditions (p-value less than 0.1 or 0.05) can possibly lead to rash conclusions. Therefore, I confirm my choice of the event window by averaging the p-values of the abnormal returns at time t over the whole sample and plotting the resulting curve. The average p-value reaches its minimum at the center of the event window (see figure 2.5).

Figure 2.4: Number of significant abnormal returns (AR) – justification of the event period

Figure 2.4 depicts the number of significant abnormal returns on the 90 respectively the 95% confidence level for each day of the event period, ranging from fifteen days before to fifteen days after the announcement. The vertical line indicates the announcement day.

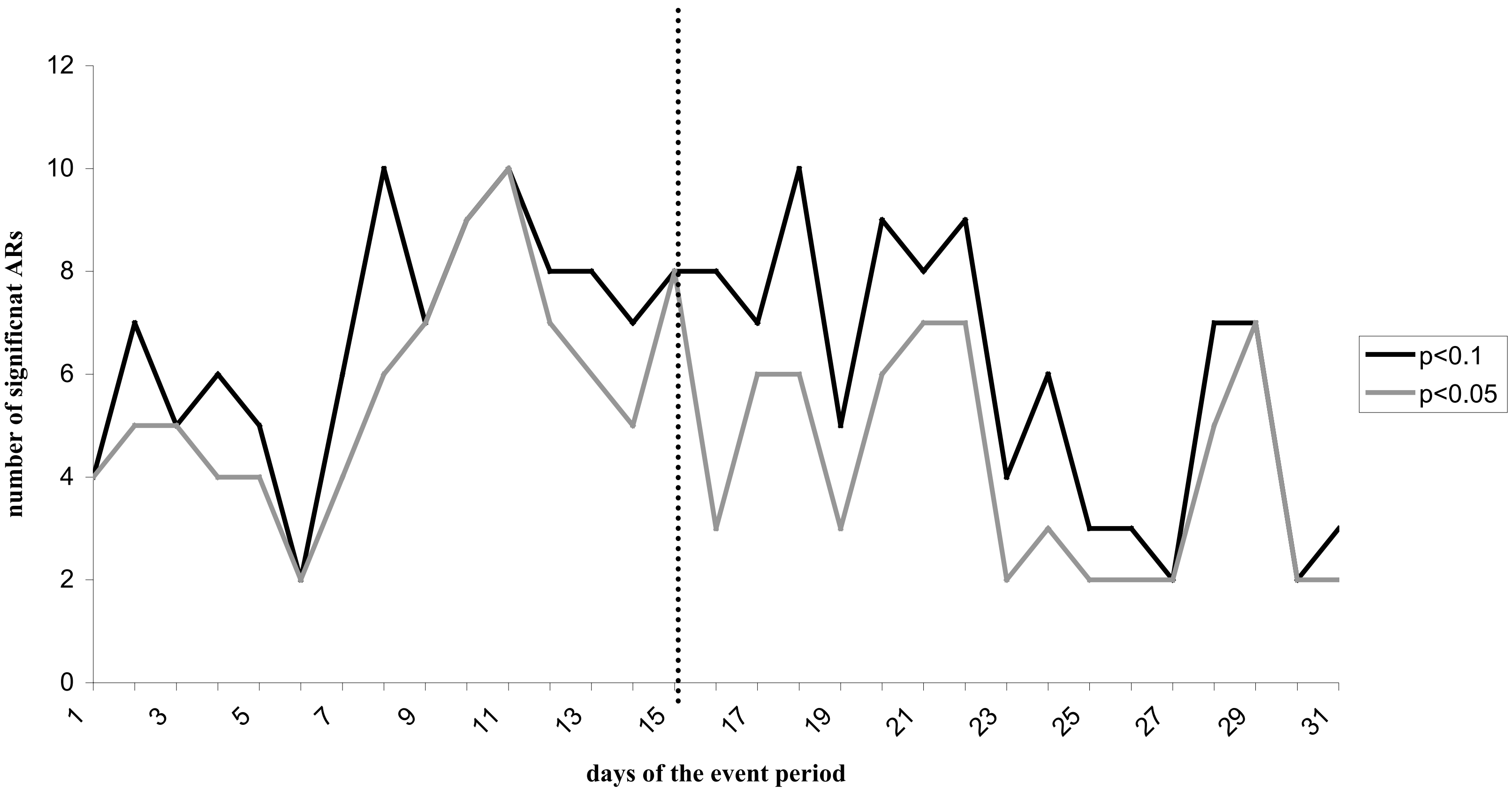
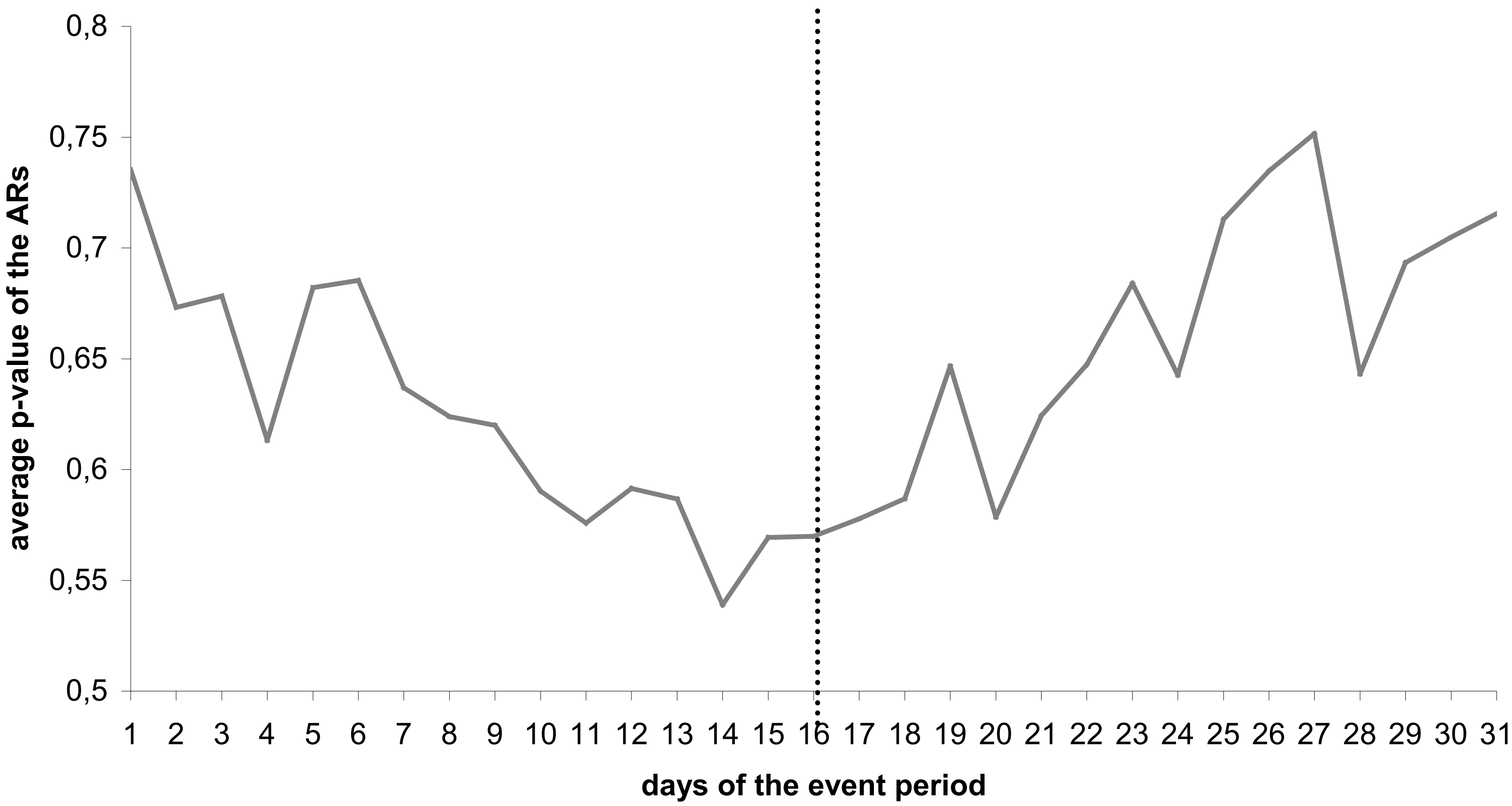


Figure 2.5: Average p-value of the abnormal returns (AR) – justification of the event period

Figure 2.5 depicts the average p-value of the abnormal returns for all cases at each day of the event period, and the vertical line illustrates the announcement day.



2.5.3 Abnormal returns and cumulated abnormal returns of the whole sample

To get a first impression about the adaptation process surrounding the merger event, I estimate the abnormal returns at time t for the whole sample of fifty companies. Therefore, I use the estimated abnormal return vector $\hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^{*}$ and aggregate over the stocks included in this vector. Following result 1 and assuming that the abnormal return at t is independent among cross-sectional units, it is straightforward to calculate an adequate test statistic. Result 3 covers this issue. I, then, turn to the aggregation over the event window and determine $\overline{C}(1;\tau_n)$; thereby, the time interval varies over which I cumulate. The test statistic is obtained from result 4. Table 2.4 summarizes the results.

All firms in the sample underwent a dramatic change of their market value during the event period. At the end of the event window, fifteen days after the announcement, the average stock price increased by 3.30 % in comparison to the expected price development, in which one believes using the CMR model. Moreover, I can draw some additional conclusions about my chosen event window. The cumulated portfolio weighted abnormal return $\overline{C}(1;\tau_n)$ exhibits significant values three day before the announcement and stays significant till the end of the period. For the purpose of illustration, figure 2.6 plots the portfolio weighted abnormal return $\bar{\varepsilon}_t^{*}$ and the values of $\overline{C}(1;\tau_n)$. Gray boxes indicate whether the realizations are significant on the 90% level of confidence.

This graph emphasizes that the adaptation process takes place within a narrow time span around the event day (t=16). Thus, one can be quite certain to capture the whole market response triggered by the merger announcement. For a deeper insight into the structure of the gains from mergers, I concentrate on two subgroups, namely acquiring and target firms.

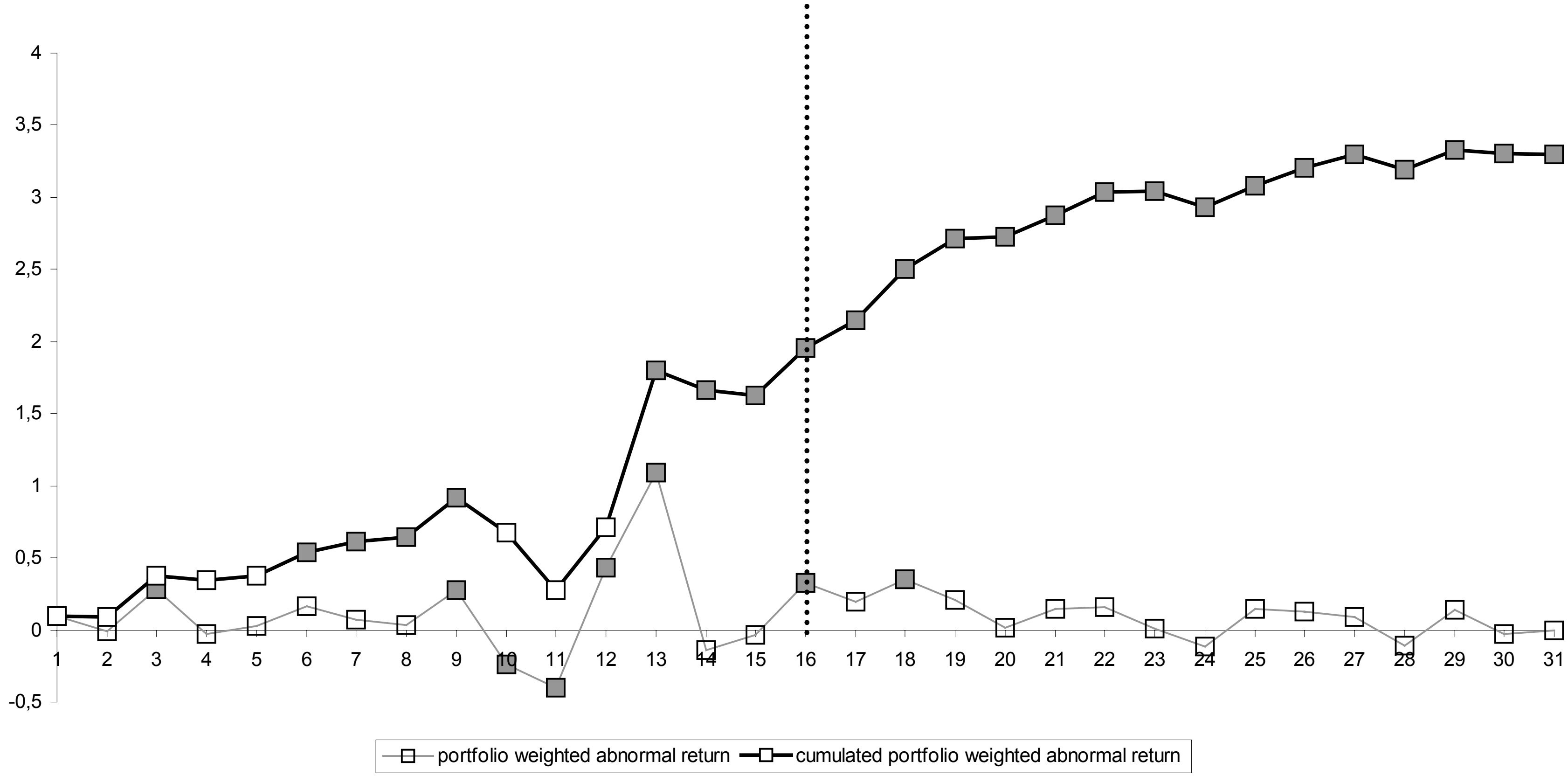
Table 2.4: Abnormal and aggregated cumulated abnormal return for the whole sample

This table contains the portfolio weighted abnormal return $\bar{\varepsilon}_t^*$ at each event day t, and the third column shows the p-value of $\bar{\varepsilon}_t^*$. The aggregation over different time intervals $\bar{C}(1;\tau_n)$ is listed and the significance is assessed by using p-values that appear in the fifth column. The event day is t=16.

$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\bar{C}(1;\tau_n)$	p-value
1	0.0982	0.458	0.0982	0.458
2	-0.0081	0.951	0.0901	0.630
3	0.2858	0.031	0.3760	0.101
4	-0.0291	0.826	0.3469	0.190
5	0.0278	0.834	0.3746	0.205
6	0.1641	0.215	0.5387	0.097
7	0.0725	0.584	0.6112	0.081
8	0.0318	0.810	0.6429	0.086
9	0.2745	0.038	0.9175	0.021
10	-0.2399	0.070	0.6775	0.105
11	-0.4025	0.002	0.2751	0.531
12	0.4347	0.001	0.7098	0.122
13	1.0913	0.000	1.8010	0.000
14	-0.1394	0.292	1.6616	0.001
15	-0.0345	0.794	1.6271	0.002
16	0.3284	0.013	1.9556	0.000
17	0.1942	0.142	2.1497	0.000
18	0.3525	0.008	2.5022	0.000
19	0.2089	0.114	2.7111	0.000
20	0.0132	0.921	2.7243	0.000
21	0.1486	0.262	2.8729	0.000
22	0.1607	0.225	3.0336	0.000
23	0.0122	0.927	3.0458	0.000
24	-0.1156	0.382	2.9302	0.000
25	0.1487	0.261	3.0788	0.000
26	0.1267	0.338	3.2055	0.000
27	0.0914	0.490	3.2970	0.000
28	-0.1078	0.415	3.1891	0.000
29	0.1402	0.289	3.3294	0.000
30	-0.0267	0.840	3.3027	0.000
31	-0.0052	0.969	3.2975	0.000

Figure 2.6: Abnormal return and cumulated abnormal return of the whole sample

Figure 2.6 contains the portfolio weighted abnormal returns $\bar{\varepsilon}_t^*$ for each day $t \in \{1,2,\ldots,31\}$ of the event window and the aggregation over increasing time intervals $\overline{C}(1;\tau_n)$. Gray boxes indicate significance on the 90% confidence level.



2.5.4 Division into two subgroups

To discuss the merger paradox, I have to evaluate whether the acquiring firms gain from mergers. Thus, the sample is divided into two subgroups, the acquiring and the acquired firms. There is almost no doubt that targets exhibit increases in their market value because the acquiring firm has to pay a premium to convince shareholders of the target to give up their ownership.[47] Furthermore, without regulatory restrictions during the pre-1914 period, an acquirer tries to buy stocks on the open market and, hence, behaves like an insider who believes that the true value exceeds the current market price. This trading behavior leads to rising stock prices.[48] In addition, after an official announcement, an acquiring firm launches an offer, which can consist of cash payment or own stocks. This offer should also have a positive impact on the stock prices of the target firm.[49] Empirical studies such as Mandelker (1974) confirmed that targets gain from mergers,[50] whereas Travlos and Papaioannou (1991) found negative abnormal returns for acquiring firms if the merger is financed by issuing new shares. Focusing on Germany, Bühner (1991) conducted a long-horizon event study[51] covering the period 1971-1985; thereby, he included 110 M&As and found that acquirers loose on average six per cent after the transaction.

The technical device of dividing the sample into subgroups is only a crude instrument to answer these questions. Because of lacking control for other stock characteristics, it is hardly possible to assess if the abnormal returns differ systematically between these two subgroups. Thus, the next step in my analysis guides me to a cross-sectional model that should have the capability to deliver clearer results. However, for a first impression, I calculate the portfolio weighted abnormal return and cumulated abnormal return, in the same fashion as before, separately for acquiring and acquired firms (see table 2.5 and 2.6).

In contrast to empirical findings for the 1970s and later periods, I find a positive cumulated abnormal return $\overline{C}(1;31)$ of about 2.27% (p-value 0.002) for acquiring firms. Thus, the merger paradox cannot be maintained for the historical period. Correspondingly, I can

[47] There are innumerable theoretical as well as empirical studies regarding the takeover premium for voting shares. For instance, a recent empirical study due to Rydqvist (1996) focused on the Swedish stock market. He regressed the relative voting premium on a variable that measures the competitiveness of the company's ownership structure. A takeover premium is only paid for voting shares. Hence, the difference between voting and non voting share prices the so called voting premium should rise, if the probability of a takeover increases or if there is a rumor about an imminent takeover.

[48] Such a behavior was possible in the pre-1914 period and is examined in chapter three.

[49] Roll (1986) argued that an overpayment is likely because the acquirer overstates the true value of the target. This assertion can be justified by an individual mistake (`arrogance´). Shleifer and Vishny (1988) stressed that typical `winners curse´ arguments could also explain this mistakes. This requires, however, that the target firm can be regarded as a common value and some potential acquirers compete against each other. Nevertheless, Jensen and Ruback (1983) did not confirm this view in their investigation.

[50] Asquith (1983), Firth (1980) showed that targets gained from mergers, whereas acquiring firms loose.

[51] I criticize the usage of event-studies for a long-term analysis in my fifth chapter and propose alternatives.

state that engaging in mergers was wealth creating for the shareholders of the acquiring firms. Moreover, the cumulated abnormal return $\overline{C}(1;31)$ reaches 5.47% (p-value 0.001) if the firm is the target of a takeover. This increase is relatively low compared to studies for later periods; hence, the premium that had to be paid to the shareholders of the target is much lower in the year 1908 than nowadays.[52] Furthermore, the adaptation process seems to differ between targets and acquiring firms. To illustrate this point, figure 2.7 depicts the cumulated portfolio weighted abnormal return for the two subgroups.

It is worthwhile mentioning that the adaptation process starts at t=13, three days before the public announcement, when the firm is target of the takeover. Beginning at t=13 the cumulated abnormal return of the target firms stays significant. In contrast, the adaptation of the fundamental value of the acquiring firm takes mainly place at the event day t=16 and after the declaration. Thus, informational motivated trading seems to play a greater role for the price process of targets compared to acquirers. Chapter three discusses this issue precisely.

Although the insights from this division into subgroups should not be understated, I am bound by the fact that I still work on an aggregate level. This means, I use portfolio weighted abnormal returns and, hence, have no chance to investigate the micro level, the respective firm and its behavior. A micro-level study should be based on a cross-sectional model that controls for company specifics. Consider that depicting the cumulated abnormal returns for single companies is less convincing because the power of event-studies declines tremendously and a distinction between mergers and other exogenous shocks is no longer reliable.[53] In chapter four, I emphasize limitations of event-studies that stem from false compositions of single events within a portfolio of securities.

[52] Even the study due to Eckbo (1986) who investigated the period from 1964 to 1983 and concentrated on merger occurring in Canada showed a higher increase in market values of target firms. Note that he detected only an average cumulated abnormal returns of about 10% which is clearly the lowest value I found in the literature.
[53] See also Morse (1984) and the discussion on the frequency of data and the sufficient number of observations.

Table 2.5: Abnormal and cumulated abnormal return for acquiring companies

Table 2.5 contains the portfolio weighted abnormal return $\bar{\varepsilon}_t^*$ at each event day t, and the third column shows the p-value of $\bar{\varepsilon}_t^*$. The aggregation over different time intervals $\bar{C}(1;\tau_n)$ is listed and the significance is assessed, using p-values that appear in the fifth column. The event day is t=16.

$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\bar{C}(1;\tau_n)$	p-value
1	0.0214	0.871	0.0214	0.871
2	0.0817	0.534	0.1030	0.580
3	0.0891	0.498	0.1921	0.399
4	0.0136	0.917	0.2058	0.434
5	0.0262	0.842	0.2320	0.430
6	0.0602	0.647	0.2922	0.364
7	0.0214	0.871	0.3136	0.367
8	0.1258	0.339	0.4394	0.237
9	0.1393	0.290	0.5786	0.142
10	-0.0170	0.897	0.5617	0.177
11	-0.1031	0.433	0.4586	0.293
12	0.1447	0.271	0.6033	0.185
13	0.2950	0.025	0.8982	0.058
14	-0.1885	0.152	0.7098	0.149
15	-0.0232	0.860	0.6866	0.178
16	0.1905	0.147	0.8771	0.095
17	0.2863	0.029	1.1634	0.032
18	0.4816	0.000	1.6450	0.003
19	0.1642	0.212	1.8092	0.002
20	0.1146	0.384	1.9238	0.001
21	-0.1229	0.350	1.8009	0.003
22	0.3418	0.009	2.1427	0.001
23	-0.0262	0.842	2.1165	0.001
24	-0.1149	0.382	2.0016	0.002
25	-0.0396	0.763	1.9620	0.003
26	0.2031	0.122	2.1651	0.001
27	-0.0266	0.840	2.1385	0.002
28	-0.0360	0.785	2.1025	0.003
29	0.1332	0.311	2.2357	0.002
30	-0.0262	0.842	2.2095	0.002
31	0.0650	0.621	2.2745	0.002

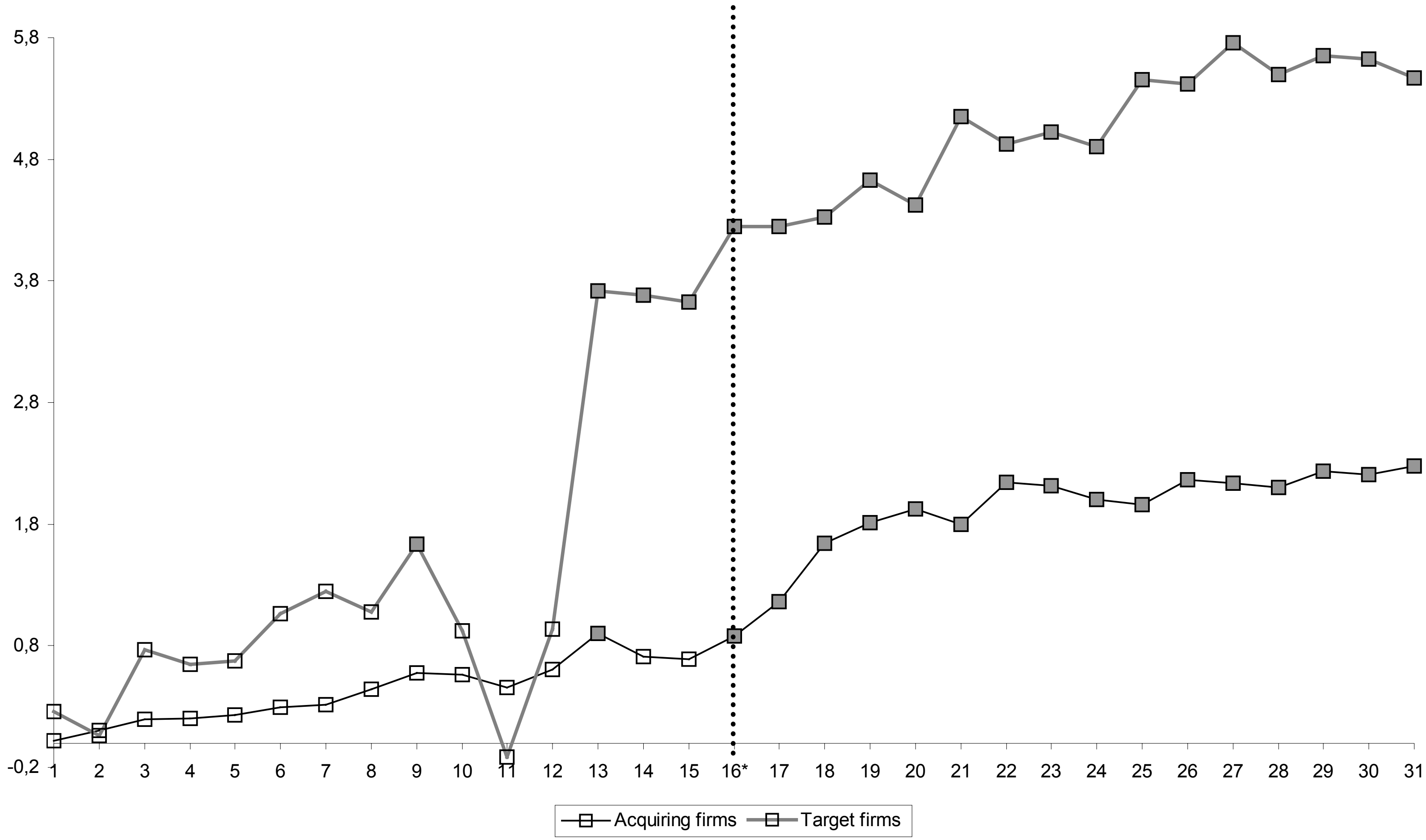
Table 2.6: Abnormal and aggregated cumulated abnormal return for target firms

Table 2.6 contains the portfolio weighted abnormal return $\bar{\varepsilon}_t^*$ at each event day t, and the third column shows the p-value of $\bar{\varepsilon}_t^*$. The aggregation over different time intervals $\bar{C}(1;\tau_n)$ is listed and the significance is assessed, using p-values that appear in the fifth column. The event day is t=16.

$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\bar{C}(1;\tau_n)$	p-value
1	0.2616	0.391	0.2616	0.391
2	-0.1990	0.514	0.0626	0.885
3	0.7039	0.021	0.7666	0.147
4	-0.1199	0.694	0.6467	0.289
5	0.0310	0.919	0.6777	0.320
6	0.3849	0.207	1.0626	0.155
7	0.1809	0.553	1.2436	0.123
8	-0.1680	0.582	1.0755	0.212
9	0.5620	0.065	1.6376	0.073
10	-0.7138	0.001	0.9238	0.338
11	-1.0386	0.001	-0.1149	0.910
12	1.0509	0.001	0.9361	0.375
13	2.7834	0.000	3.7195	0.001
14	-0.0351	0.908	3.6844	0.001
15	-0.0586	0.848	3.6258	0.002
16	0.6215	0.042	4.2473	0.001
17	-0.0016	0.996	4.2457	0.001
18	0.0780	0.798	4.3237	0.001
19	0.3041	0.319	4.6278	0.001
20	-0.2022	0.507	4.4256	0.001
21	0.7254	0.017	5.1510	0.000
22	-0.2242	0.462	4.9268	0.001
23	0.0938	0.758	5.0206	0.001
24	-0.1172	0.701	4.9034	0.001
25	0.5487	0.072	5.4521	0.000
26	-0.0357	0.907	5.4164	0.001
27	0.3423	0.262	5.7587	0.000
28	-0.2605	0.393	5.4982	0.001
29	0.1552	0.611	5.6534	0.001
30	-0.0276	0.928	5.6258	0.001
31	-0.1544	0.613	5.4714	0.001

Figure 2.7: Cumulated abnormal return of acquiring and target firms

Figure 2.7 plots the aggregated cumulated abnormal return for increasing intervals starting at t=1 and ranging till t=31; thereby, I divide between acquiring and target firms.



2.6 Cross sectional model

Working on a firm specific level, I can overcome the limits of analyzing subgroups and shed some light on the driving forces for successful mergers. By controlling stock specific characteristics, the partial impact of belonging to the group of acquirers on the change in market values can be measured. Hence, a precise statement regarding the existence of the merger paradox is possible.

The cumulated abnormal return of firm i $\overline{C}(1;31)$ as a measure of success is, thereby, the dependent variable. This cumulated abnormal return is derived as mentioned in result 2. Furthermore, I introduce some potential explanatory variables. Besides the dummy variable $Target_i$ which takes the value one, if the firm is a target, and zero otherwise, I include firm size, former dividend payments, line of business, the way the merger is financed, the age of the firm and a dummy which takes the value one if the management of the target is replaced after the merger. By construction of my sample in which I include mergers that fail to be carried out, I can use this additional information to explain the cumulated abnormal returns. Reading the `Berliner Börsenzeitung´ provides information about the failure to reach the necessary majority in the shareholder gathering, or if the advisory board respectively legislative obstacles prevent a takeover. In addition, I also know the exact day on which the disapproval and other circumstances that make the merger unlikely are publicly declared.

Before discussing the model structure, I should mention the expected impact of these explanatory variables, the economic intuition, and theoretical justification to include these characteristics. As seen above, the dummy indicating targets could possess an influence on cumulated abnormal returns. If a remarkable discrepancy in cumulated abnormal returns between targets and acquirers was detected, the results obtained from my former event-study regarding the merger paradox might have to be rethought.

I measure firm size by calculating the market capitalization of firm i. Therefore, I extract from the `Handbuch der Deutschen-Aktien-Gesellschaften´ the par value of the issued shares and the nominal capital to calculate the number of outstanding shares. Then, the number of outstanding shares is multiplied with the actual share price at the beginning of the event period. There are several reasons, why one should control for firm size. If a relatively big company takes control over a small target firm, I should not expect tremendous wealth effects for the bigger company. This is caused by the fact that the newly purchased small firm contributes only a tiny piece to future earnings of the acquiring firm. Moreover, the access to capital for financing takeovers is easier and cheaper for big firms. Tilly (1982), for instance, claimed that market imperfections are mainly responsible for the connection of firm size and access to capital. In addition, size might also serve as a weak proxy for economies of scale in an industry.[54]

He also pointed out that "(…) the costs of raising funds on the main capital markets in Germany varied inversely with the age (…) of industrial corporations."[55] According to this statement and the empirical evidence provided by Tilly (1980), I decide to insert age of the company in my model. Moreover, the age of the firm serves as a proxy for its capability to survive and its experience.

The dividend payments of three successive years give a hint with regard to the profitability of the firm. Consequently, the annual growth rate of dividend payments is my measure for changes in profitability before the merger.[56] A very profitable firm can be watered down after a merger; thus, the market could punish a merger announcement.

The line of business categories indicate industry specific factors that could influence the success of mergers. For instance, acquiring banks justify their decision by emphasizing the need to expand in order to serve new regional markets and to reduce credit risks by diversification.

Travlos and Papaioannou (1991) attached some importance to the way of financing a takeover; thereby, the theoretical consideration is based on Leland, Pyle (1977), Myers and Majluf (1984). They suggested that under asymmetric information, managers of the acquiring firm prefer cash payment if their share is undervalued, whereas exchange of common stocks is favored if their firm is overvalued. Consequently, the signaling models predicted that cash offers can be regarded as good signal and should, therefore, yield positive abnormal returns for the respective company. Furthermore, Travlos and Papaioannou (1991) calculated changes in financial leverage; thus, they controlled for capital structure effects after acquisitions.[57] According to lacking data quality and availability, I can only make a rough distinction between cash payment or a transfer of shares. Henceforth, determining the financial leverage is hardly reliable for my investigation period. Obviously, this limits my analysis – but a further refinement would inevitably reduce the number of observations.

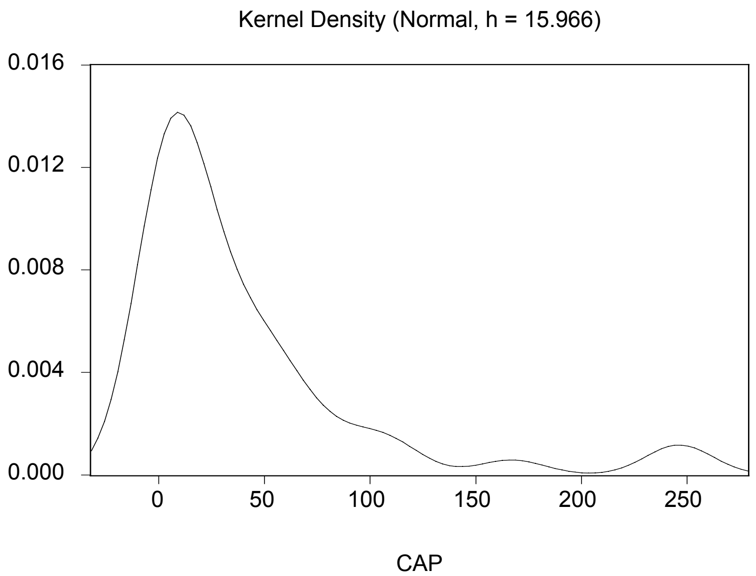
---

[54] I thoroughly discuss the literature regarding the alleged advantages of size in chapter one.
[55] See Tilly (1982, p. 645).
[56] Note that during the pre-World-War I period, a strong interrelation between earnings and dividends existed. The smoothing of short-term fluctuations in earnings by following a dividend policy was less common. Nevertheless, reported earnings itself were often manipulated. Accordingly, dividend payments are more reliable estimates.
[57] This enables to distinguish between a method of payment and a capital structure effect. A higher leverage (relative increase of loan capital) yield higher abnormal returns because it reduces the free cash flow (see Jensen, 1986).

If I trust in the argument, provided at first by Manne (1965), that the market for corporate control and the related principal agent problem is essential for explaining takeover activities the replacement of an incompetent management should yield positive returns.

Before starting with a very simple cross-sectional model, it is worthwhile to look how the variables are distributed. The market cap, age of the firm and the cumulated abnormal return are not symmetrically distributed. This can be represented by the Kernel density, which is simply speaking the continuous alternative to histograms. Figure 2.8 plots the Kernel density for market capitalization.

Figure 2.8: Kernel density for the variable market capitalization



The distributions of the other two variables look pretty much the same; thus, I skip these illustrations. Using such a lopsided explanatory variable could result in a violation of CLR assumptions, especially, that the linear structure of the model is correctly specified. To offset this problem, I transform the market capitalization and the age of the firm by taking the natural logarithm. Obviously, taking the natural logarithm of the cumulated abnormal return would lead to a loss of observations because some returns are negative. Thus, I construct an index ranging from zero to 100 and take the natural logarithm of these indexed values.

$$\hat{C}_i^{index}(1;31) = \log\left[\frac{\hat{C}_i(1;31) - Min_i\hat{\mathbf{C}}(\mathbf{1;31})}{Max_i\hat{\mathbf{C}}(\mathbf{1;31}) - Min_i\hat{\mathbf{C}}(\mathbf{1;31})}\cdot 100\right] \tag{2.16}$$

I use, therefore, the following simple model structure for getting a first impression. Note that I include only two industry specific dummies, namely for the banking industry and the mining sector. Refining this division further could cause severe problems. To illustrate this point, consider the brewery industry for which I have only one observation, `Lindener

Aktienbrauerei´. Thus, a dummy variable defined for breweries explains the deviation of only one observation in comparison to a reference group. That does not lead to reliable results and interpretations. The following regression equation is, hence, my basic model.

$$\hat{C}_i^{index} = \beta_0 + \beta_1 \log(cap_i) + \beta_2 \log(age_i) + \beta_3 Success_i + \beta_4 Change_i + \tag{2.17}$$

$$\beta_5 Cash_i + \beta_6 DivGrowth_i + \beta_7 Bank_i + \beta_8 Mining_i + \beta_9 \mathrm{Target}_i + u_i$$

Where: β_0… Intercept

- $\mathrm{Log}(cap_i)$… Natural logarithm of the market value of stock i
- $\mathrm{Log}(age_i)$… Age of firm i
- $Success_i$… Dummy that takes value one if the merger is executed
- $Change_i$… Dummy that takes value one if the management is replaced
- $Cash_i$… Dummy that takes value one if the merger is financed by cash.
- $DivGrowth_i$… Annual growth rate of dividend payments (1906 to 1908)
- $Bank_i$… Dummy that takes value one if firm i belongs to the banking industry
- $Mining_i$… Dummy that takes value one if firm i belongs to the mining industry
- $Target_i$… Dummy that takes value one if firm i is the target of a merger
- u_i… is the disturbance term

Carrying out regression (2.17), I receive the results, shown in table 2.7. Moreover, table 2.7 contains the p-values of a White-test and a F-test, testing for heteroscedasticity and the explanatory power of the whole model. Note that interaction terms between the dummy variable for target firms and other explanatory variables does not show any significance.[58]

Arguing on the 10% level of significance, belonging to the mining or banking industry and showing high growth rates of dividend payments[59] over the last three years are important factors that yield higher cumulated abnormal returns. In contrast to my considerations, old firms exhibit systematically lower cumulated abnormal returns. Maybe the age of a firm is not a good proxy for experience or access to capital. But before drawing false conclusions, one should test for misspecification problems that could bias the OLS estimates.

[58] Based on Travlos and Papaioannou (1991), one could suggest that interactions between the dummy for target firms and the way of financing a merger would be relevant for explaining cumulated abnormal returns. Recall that the theoretical justification provided by the signaling models referred only to the performance of acquiring firms. Nevertheless, the p-value of the coefficient of the interaction term reaches only 0.655; thus, I cannot confirm any differences between targets and acquiring firms if the merger is financed by cash payment or common stock exchange.

[59] This is not line with my expectations – but due to omitted variable problems, one should not overstate this preliminary result.

Table 2.7: OLS estimation of regression equation (2.7)

Explanatory variable	Coefficients	p-value
Intercept	4.5720	0.000
$Log(cap_i)$	-0.0494	0.592
$Log(age_i)$	-0.4532	0.009
$Success_i$	-0.6415	0.266
$Change_i$	-0.4284	0.916
$Cash_i$	-0.1943	0.937
$DivGrowth_i$	0.5855	0.089
$Bank_i$	0.4775	0.098
$Mining_i$	0.6891	0.074
$Target_i$	-0.2004	0.435
Number of Observations N	45	-
Adjusted R^2	0.14	-
F-Test (p-value)	1.80	0.104
White-Test NR^2 (p-value)	43.68	0.177

Recall that by definition a higher normal return yield smaller abnormal returns. Hence, one may wonder if the normal return should be included in my model. A log likelihood ratio test[60] tells us that I should control for the estimated mean vector μ and, hence, include this variable. The log likelihood ratio reaches 19.25 (p-value 0.000), which does not permit doubts about the specification problem. After including the estimated mean μ_i of stock i, I obtain model one depicted in table 2.8.

$$\hat{C}_i^{index} = \beta_0 + \beta_1 \log(cap_i) + \beta_2 \log(age_i) + \beta_3 Success_i + \beta_4 Change_i + \beta_5 Cash_i + \beta_6 DivGrowth_i + \beta_7 Bank_i + \beta_8 Mining_i + \beta_9 Target_i + \beta_{10} Mean_i + u_i \tag{2.18}$$

I also calculate the reduced form (2.19); hence, I regress the estimated mean vector μ on the exogenous explanatory variables of (2.17).[61] Model two in table 2.8 shows the results of this procedure and uncovers that belonging to the mining industry has a negative effect on estimated means, whereas older companies exhibit higher estimated mean returns.

$$Mean_i = \gamma_0 + \gamma_1 \log(cap_i) + \gamma_2 \log(age_i) + \gamma_3 Success_i + \gamma_4 Change_i + \gamma_5 Cash_i + \gamma_6 DivGrowth_i + \gamma_7 Bank_i + \gamma_8 Mining_i + \gamma_9 Target_i + v_i \tag{2.19}$$

What do these results tell us? The interpretation is not straightforward because the significant coefficient of μ_i in model one raises several problems. By construction, the cumulated abnormal returns, even after the transformation, depend on the estimated mean vector.

[60] The restricted model is regression (2.17), and the unrestricted model includes, besides the set of exogenous variables from regression (2.17), the estimated mean as additional explanatory variable.

[61] This reduced form is just motivated by econometric considerations to control for every possible indirect effects. After simplifying the model, I can easily interpret the outcomes in an economical sense. For instance, some industries, like the mining sector, might exhibit a pronounced drift during the estimation period due to exogenous shocks.

Defining the abnormal return (2.9) as difference of the observed return in the event period minus the estimated mean, I should suggest that the higher the estimated mean of a stock i the smaller the abnormal return and, hence, a negative coefficient is expected and observed.

In addition, I should consider a causality problem. This causality problem is the main reason, to include μ_i into regression (2.17). To illustrate this point consider the following. If I neglect μ_i in regression (2.17) and detect a significant influence of an explanatory variable on the cumulated abnormal return, this impact can stem from two sources. First, the variable is responsible for changes in the cumulated abnormal return of stock i during the event period. Second, the explanatory variable has an impact on the estimated mean μ_i; hence, by definition, if μ_i changes, the cumulated abnormal return is changed. This indirect influence represents the causality problem. An explanatory variable that possesses an impact on μ_i, which is determined in the estimation period, has no causal relation to the cumulated abnormal return, determined in the event period.

Table 2.8: Outcomes of regression (2.18) and regression (2.19)

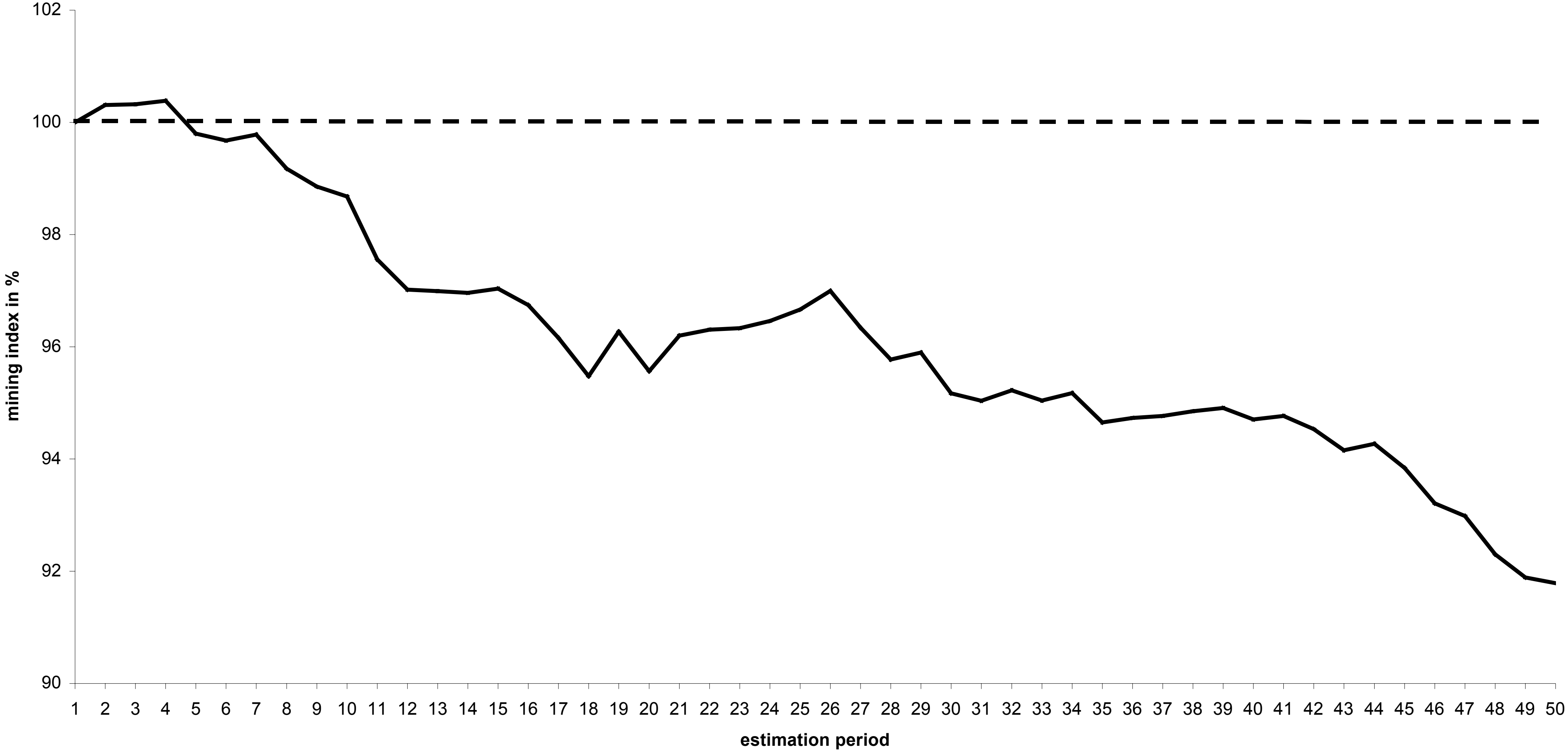
Table 2.8 shows the OLS output of regression (2.18) in column two called model one. The columns three contains regression (2.19) with the estimated mean vector as dependent variable. P-values are set in parentheses.

Explanatory variable	**Model 1** Dependent Variable $\hat{C}(1;31)$	**Model 2** Dependent Variable $\hat{\mu}$
Intercept	3.3960 (0.000)	-0.4026 (0.007)
Log(cap_i)	-0.0107 (0.888)	0.1428 (0.420)
Log(age_i)	-0.1710 (0.260)	0.0967 (0.005)
$Success_i$	-0.4113 (0.385)	0.0736 (0.511)
$Change_i$	0.0188 (0.955)	0.0230 (0.775)
$Cash_i$	-0.1878 (0.363)	-0.0557 (0.252)
$DivGrowth_i$	0.4347 (0.125)	-0.0532 (0.423)
$Bank_i$	0.4321 (0.070)	-0.0149 (0.790)
$Mining_i$	0.2143 (0.516)	-0.1595 (0.037)
$Target_i$	-0.1137 (0.590)	0.0274 (0.584)
$Mean_i$	-2.9462 (0.000)	-
Number of Observations	45	45
Adjusted R^2	0.42	0.21
F-Test (p-value)	4.22 (0.001)	2.35 (0.033)
White Test NR^2 (p-value)	45.00 (0.348)	43.64 (0.179)

Consider that the dummy for the mining industry has a negative impact (p-value: 0.037) on the estimated mean return. This means that stocks of the mining industry decline remarkably during the estimation period (see figure 2.9). Exogenous shocks, for instance a decline in raw material prices, that are not taken into account by my model are responsible for the bad performance of mining companies. If I control for this effect on the normal return, a direct impact is not observable. The p-value of the dummy $mining_i$ reaches only 0.516 in model one of table 2.8. Furthermore, old companies affect the estimated mean in a negative manner (p-value: 0.005) – but the age has no direct impact on the success of mergers (p-value: 0.260). Controlling for exogenous influences on the determined normal return enables to evaluate the influence of the dummy $bank_i$. The banking industry exhibits significantly higher cumulated abnormal returns; thus, mergers among banks are more successful.

Note that thus far model one and two (see table 2.8) are estimated using system OLS. Obviously, this estimation procedure stays unbiased as long as exogeneity conditions hold; thereby, I have to ensure that for both equations the explanatory variables are uncorrelated with the error term of the respective equation. Accordingly, I conduct Hausman procedures to test whether the estimated mean is endogenous. Considering the reduced form (see model two in table 2.8), one can argue that $log(age_i)$ and the dummy for the mining industry possess a partial impact on the estimated mean return – but they do not influence the ultimate dependent variable, namely the cumulated abnormal return. Thus, using these two explanatory variables as instruments for the estimated mean return seems to be appropriate. If I use both variables as instruments, the coefficient's p-value of the residual of the reduced form reaches 0.205. If only $log(age_i)$ is used as instrument the p-value of the residual is 0.260, and if only the dummy $mining_i$ serves as instrument the p-value is 0.516. So the exogenity assumption is maintained, and OLS estimation is still consistent. Nevertheless, I carry out a simultaneous equation procedure to show that my empirical findings do not depend on the chosen procedure.

Figure 2.9: Declining stock prices in the mining industry during the estimation period



Using a three-stage least[62] squares procedure requires the definition of a set of instrumental variables that have to be sufficiently large to identify the linear equation system. For that purpose, I orient toward my results for model two (see table 2.8), which indicate that *log(age$_i$)* affects the mean significantly. Therefore, this exogenous explanatory variable serves as an instrument for the mean vector. Thus, *log(age$_i$)* is canceled in the first equation and used as instrument in the second equation only. This leads to the following system. Table 2.10 provides the outcomes of the three-stage least squares estimation.

$$\hat{C}_i^{index} = \beta_0 + \beta_1 \log(cap_i) + \beta_2 Success_i + \beta_3 Change_i + \beta_4 Cash_i + \beta_5 DivGrowth_i + \beta_6 Bank_i + \beta_7 Mining_i + \beta_8 \mathrm{Target}_i + \beta_9 \mu_i + u_i \tag{2.20}$$

$$\mu_i = \delta_0 + \delta_1 \log(cap_i) + \delta_2 \log(age_i) + \delta_3 Success_i + \delta_4 Change_i + \delta_5 Cash_i + \delta_6 DivGrowth_i + \delta_7 Bank_i + \delta_8 Mining_i + \delta_9 \mathrm{Target}_i + v_i$$

Following the rule, from general to specific, I reduce the model to detect the essential influential factors for the cumulated abnormal return. After executing an F-test[63] that imposes linear restrictions on model (2.20) and carrying out a specification test that compares the restricted with the unrestricted model,[64] I obtain the model presented in table 2.11.

$$\hat{C}_i^{index} = \beta_0 + \beta_6 Bank_i + \beta_8 Mean_i + u_i \tag{2.21}$$

$$Mean_i = \delta_0 + \delta_2 \log(age_i) + \delta_8 Mining_i + v_i$$

Using the ten per cent level of significance, I can argue that the banking industry exhibits larger cumulated abnormal returns. Moreover, the crucial point that the dummy *bank$_i$* could have an influence on the estimated mean μ_i is ruled out by the simultaneous equation model. Hence, I can avoid the problem of misspecification as well as doubts about the causal relation between the explanatory variables and the cumulated abnormal return. Furthermore, the exogenous explanatory variables *mining$_i$*, and *log(age$_i$)* influence significantly the estimated mean μ_i – but do not affect the cumulated abnormal return, determined in the event period, in a causal manner. The former empirical findings applying system OLS are confirmed by this simultaneous equation procedure.

---

[62] Note that a equation by equation two-stage least squares estimator is algebraically identical because the each equation is just identified. Wooldridge (2002) provided the proof.

[63] All coefficients are supposed to be equal to zero apart from β_6, β_8, δ_2, and δ_8. This joint hypothesis cannot be rejected because the F-test statistic reaches 0.77 (p-value 0.701).

[64] The likelihood-ratio test shows a test statistic of 10.0 (p-value 0.761); thus, the null hypothesis that the model is correctly specified is not rejected.

Table 2.10: Outcomes of the three-stage least squares estimation of model (2.20)

P-values appear in parentheses.

Explanatory variable	Equation1: Dependent Variable $\hat{C}_i^{index}$	Equation2: Dependent Variable Mean$_i$ μ_i
Intercept	2.6831 (0.000)	-0.3992 (0.002)
Log(cap_i)	0.0127 (0.870)	0.0131 (0.419)
Log(age_i)	-	0.0958 (0.001)
Success$_i$	-0.2718 (0.545)	0.0781 (0.438)
Change$_i$	0.0561 (0.857)	0.0209 (0.771)
Cash$_i$	-0.2898 (0.144)	-0.0571 (0.189)
DivGrowth$_i$	0.3434 (0.200)	-0.0512 (0.389)
Bank$_i$	0.4045 (0.068)	-0.0154 (0.757)
Mining$_i$	-0.0735 (0.847)	-0.1612 (0.017)
Target$_i$	-0.0611 (0.759)	0.0294 (0.513)
Mean$_i$	-4.7322 (0.001)	-
Number of Observations	45	45
"Adjusted R^2"	0.47	0.36
F-Test (p-value)	2.96 (0.005)	2.83 (0.007)

Table 2.11: Results of model (2.21) after excluding negligible[65] explanatory variables

P-values are set in parentheses.

Explanatory variable	Equation1: Dependent Variable $\hat{C}_i^{index}$	Equation2: Dependent Variable Mean$_i$ μ_i
Intercept	2.2702 (0.000)	-0.3022 (0.001)
Log(cap_i)	-	-
Log(age_i)	-	0.9290 (0.001)
Success$_i$	-	-
Change$_i$	-	-
Cash$_i$	-	-
DivGrowth$_i$	-	-
Bank$_i$	0.4205 (0.022)	-
Mining$_i$	-	-0.1299 (0.028)
Target$_i$	-	-
Mean$_i$	-4.4033 (0.000)	-
Number of Observations	45	45
"Adjusted R^2"	0.45	0.29
F-Test (p-value)	8.46 (0.000)	9.10 (0.000)

[65] Note that multicollinearity can be ruled out for my data set because the correlation coefficient between two explanatory variables reaches on maximum an absolute value of 0.4660 (correlation between estimated *mean$_i$* and *log(age$_i$)*). This value is obviously far away from 0.85 which is often used as critical boundary based on a rule of thumb (see Kennedy, 1998). Moreover, auxiliary regressions do not point into the direction of multicollinearity.

2.7 Conclusion

As main result, I should stress the high degree of market efficiency in the year 1908 in Germany because a sufficient velocity of information flows is a prerequisite for working with the event study approach. The adaptation process that ends when the merger announcement is fully reflected in the market prices is timely very close to the public declaration. Consequently, the market responds very quickly. More precisely, the adaptation process starts about three days prior to the release of information if I focus on target firms; therefore, informational motivated trading is apparent. Chapter three puts some emphasis on informed trading. My empirical finding also provokes additional doubts whether using weekly or monthly returns is reliable because markets reacted faster than assumed by Banerjee, Eckhard (2001), Leeth, Borg (1994, 2000), and Borg et al. (1989).

Distinguishing between acquiring and target firms and calculating the group specific cumulated aggregated abnormal return uncovers that the merger paradox can be rejected for the year 1908 in Germany because acquiring companies exhibit an increase in their stock prices of about 2.27%. Target firms exhibit an upsurge of their market values by 5.47%. Due to the fact that rejecting the merger paradox under weak economic circumstances like in 1908 is a stronger result compared to a bullish period like around 1906 (see figure 2.1), one can infer that mergers were market value increasing during the pre-World-War I period.

Note that Leeth and Borg (1994) who covered a similar period, namely 1905 to 1930, did not reject the merger paradox for their 191 mergers occurring among U.S. manufacturing companies. They found considerable gains prior to the merger for acquirers; however, this increase in market values was outweighed by a pronounced decline in share prices after the transaction. This finding might stem from their chosen period because in the 1920s and, especially, after the stock market crash 1929 the regulation regarding mergers became more important. Nevertheless, the Security Exchange Act[66] was introduced five years after the stock market crash; thus, the prohibition of insider trading and additional restrictions were established after the end of the investigation period of Leeth and Borg (1994).

Noteworthy, the increase in market values of targets is relatively low compared to studies for later periods like Eckbo's (1986) investigation who considered mergers from 1964 to 1983. He found an increase in market capitalization of about 10% which can be regarded as the lower limit compared to other studies for the second phase of globalization. Henceforth, one should wonder why companies could acquire competitors without paying considerable premiums. One answer is the possibility to conduct insider trading and to buy shares of target

[66] Leland (1992) discussed the impact of the Security Exchange Act in this introduction.

companies prior to an imminent merger. The third chapter concentrates on this issue. Furthermore, the discussion on the voting premium (see Rydqvist, 1996) can also contribute to clarify my finding. Although due to lacking data on non-voting shares, determining the voting premium for my historical sample is not possible. However, I found `narrative´ evidence in newspapers that acquiring firms bargained with principal shareholders prior to merger announcements. On 30th March 1908, page 2 of the evening issue, the `Rheinisch-Westfälische Disconto-Gesellschaft´ indicated its willingness to acquire the `Krefelder Bank´; thereby, the principal shareholders declared to give up their voting rights before the official offer. Accordingly, a voting premium for the rest of the outstanding shares was not required.

Besides these general findings, I develop a simultaneous equation model that controls for the impact of the estimation period on the cumulated abnormal returns, which I want to explain. Thus, I can conclude that the banking industry exhibits cumulated abnormal returns that are above average. This relationship is causal in the sense that belonging to the banking industry has not an effect on the estimated normal return of the respective firm. Other potential factors such as the firm size, measured by the market capitalization, do not affect the cumulated abnormal return. Furthermore, the dummy variable for target firms is never significant. Therefore, if one controls for other stock characteristics, a partial effect of being target of a transaction cannot be detected. This gives an additional hint that the merger paradox is not present.

Obviously, the market responded very positively if banks undertook mergers compared to other industries. I observe that a huge portion of the merger activities was due to the banking industry, and the market supported this strategy. But what were the reasons for this strong merger activity? Why did the market reward this behavior? From my data source, the `Berliner Börsenzeitung´, I get a clue about the motives of mergers in the banking industry. Every merger was justified by the need to enlarge the company and to get access to additional regions and markets. This geographical diversification seemed to be appropriate to reduce risk, especially in the credit business, because the bank depended less on regional-specific economic shocks. Moreover, the acquired, usually small, banks were carried on as a new branch of the acquiring bank. To keep the customers and to stay trustworthy, the old management of the target firm became – in almost all cases – the branch management. Despite this `narrative´ evidence, a rigorous empirical proof of this argumentation is not possible with my data set because counterexamples, namely banks that use mergers not for the purpose of expansion, are lacking. Thus, analyzing the motives of mergers in different lines of business should be done by comparing different periods or in cross-country studies; hence, I leave this interesting topic for future research.

In addition, I make some very restrictive assumptions to derive my test statistics based on a modified CMR model. These assumptions, such as no cross-correlation – despite common in the literature[67] using share time series models – have possibly a great influence on my results. I discuss these problems and further limitations of event-study analysis in an additional chapter.

[67] See, for instance, Dyckman et al. (1984).

3. Disclosure of mergers without regulatory restrictions: Who gains from mergers?

3.1 Extended abstract

Who gains from mergers? I concentrate on insiders and outsiders by investigating the adaptation process of stock prices around public merger announcements. The way of disclosure is essential. If firms hide information, they will hurt outsiders. Hiding information does not yield higher cumulated abnormal returns – but the higher the expected gains from mergers the higher the incentive to hide information. Hence, it should be worthwhile to restrict insider trading by forcing firms to uncover mergers. In contrast to the year 1908, pre-merger gains in the year 2000 are due to irrational speculation and not to insider-trading.

3.2 Introduction

For the year 1908 in the German case, the merger paradox that stock prices of acquiring firms decline considerably after the merger is made public was rejected[68]– but one central question is still not answered: Which shareholders gained from increases in market values? Focusing on two types, namely insiders and outsiders, this chapter tries to answer this question; thereby, the so called run-ups serve as a measure for insider gains. Run-ups are changes in share prices triggered by an impending merger announcement. Because the merger is not yet public information, significant changes prior to the public announcement are a hint for insider-trading. If market participants have only access to public information like newspapers, they belong to the group of outsiders. In contrast, insiders hold private information; hence, they already know that a firm will involve in merger activities. This superior knowledge leads to trading activities of insiders before newspapers make mergers public. Through this insider trading, the private information is conveyed; thus, the market price is significantly influenced already before the public release of mergers. Keown and Pinkerton (1981) used this pre-adjustments to uncover insider activities around revealed mergers occurring in a four-years-period starting 1975. Banerjee and Eckhard (2001) provided evidence for insider-trading in the period from 1896 to 1903, which is known as the first merger wave. Both studies concentrate on the US case. In contrast to Keown and Pinkerton (1981), Banerjee and Eckhard (2001) used weekly returns which can cause problems with regard to the power of event-study methods. They collected 56 companies from which 37 announced their willingness to merge in time and 19 tried to hide information. Unfortunately, using weekly

[68] In chapter two, I showed that the market values of merging firms increased in the year 1908.

instead of daily returns makes it more difficult to detect abnormal stock price movements caused by an announcement. The longer the chosen sampling interval the more cross-sectional units are needed to maintain a high power of the test statistics. Morse (1984) presented a precise analysis on the usage of daily versus weekly respectively monthly returns. Therefore, I collected daily returns for the samples of the years 1908 and 2000. In addition, this enables to compare the scale of insider-trading between two distinct time periods reliably because both studies are based on the same frequency of data. Of course, using intra-daily returns like recommended by Barclay and Litzenberger (1988) would further improve the power of event-studies – but it is impossible for historical time periods.

Lacking regulatory restrictions are responsible for the appearance of two different forms of disclosure. Some firms announce mergers after these mergers have already been executed, and others declare their desire to merge before the transfer of assets. This firm behavior is only observable in the sample of the year 1908 before the legal framework comes into force that does not allow to wait for disclosing mergers.

Thus, one consideration is to assess whether the way of disclosure influences the gains respectively losses of insiders and outsiders. This should help to understand why some firms choose to disclose new information and others try to hide it.

Are there incentives to disclose information? I test the hypothesis that revealing information facilitates to raise up money. If issuing new shares finances the merger, the firm needs to keep trustworthy to attract outsiders as a source of financing.

In addition, drawing a sample of merger announcements occurring in the year 2000 in Germany makes a comparison of the gains of insiders between time periods with or without regulation possible. Maybe, this sheds some light on the impact of regulations on insider activities and the ability of legislative restrictions to protect outsiders from insider trading.

If regulation can reduce insider activities, is it desirable from a normative point of view to impose these restrictions? The two most important theoretical contributions are Leland (1992), Easley and O'Hara (1992). The following pros and cons of insider trading are often discussed in literature. Brudney (1979) and others argued that insider trading is unfair – but it is an open question what the expression unfair means. Especially, if becoming an insider is costly, one should expect that holding private information yields extraordinary profits to compensate for the costs of being better informed.[69] If outsiders regard insider trading as unfair, their incentives to invest are lower. Moreover, the liquidity declines in the

[69] See also the discussion about the so called Grossman – Stiglitz (1980) paradox.

presence of insider trading[70] and current stock prices are more volatile. Less liquidity and higher volatility hurt outsiders who initiate trades because of an urgent demand for liquidity.[71] But insider trading increases the velocity with which new information is reflected in current market prices. Hence, the market works with a higher degree of informational efficiency. Consequently, from a normative point of view, it remains unclear whether insider trading should be prohibited. Is it possible to obtain an empirical justification for the need of regulation?

I try to give answers to these questions by organizing my chapter as follows. First, I describe the legislative framework in Germany. Second, the method of sampling is presented followed by a brief discussion of the event-study approach; thereby, I refer to my former results thoroughly discussed in chapter two. Third, I discuss my empirical findings whether insider-trading leads to a redistribution of total efficiency gains due to mergers from less informed to better informed traders. Fourth, the differences regarding insider-trading after 92 years of regulation are highlighted. Fifth, to obtain a clear normative statement, I exclude the possibility that hiding information influences the success of a merger as measured by the total cumulated abnormal return.

This cumulated abnormal return serves as a measure of the efficiency gains from mergers.[72] So I end up with the distributive effect of undisclosed mergers, which enables to favor regulations like the ad-hoc-publication requirement. Sixth, I try to figure out whether there are incentives to disclose voluntarily. Noteworthy, I found that greed for money is the only motive that prevents a management from revealing information. The way of disclosure is not linked to the way of financing a merger. Some concluding remarks underline the main results and outline further research topics for the future.

---

[70] See Kling (2002b) in which I detected decreasing liquidity around earnings announcements that stems from an increase in information asymmetry on the market.

[71] Obviously, this argumentation roles out the possibility of noise trading as discussed in Black (1986); thereby, an incentive to trade is imposed by irrational trading signals.

[72] In the literature, there exists a controversial position called redistribution theory. This theory suggests that shareholders gain from mergers by higher stock prices and, thus, higher cumulated abnormal returns, whereas other stakeholders like bondholders loose from mergers. Hence, the cumulated abnormal return is not an appropriate measure of the efficiency gains from mergers. A unambiguous empirical evidence for this theory is lacking up to this point. Jarrell et al. (1988) provided a discussion of this issue, and ,in chapter four, I dedicate a section to justify why an econometric study is not promising for the pre-1914 period.

3.3 Historical Background – Insider Regulation in 1908 and 2000

3.3.1 Legal Framework in the year 1908 in Germany

The exchange law (`Börsengesetz´) of 1896[73] provided the legal framework for the time period 1896 to 1914, when the first world war was started. Note that the exchange law was slightly modified, and, thus, a new version was established at April 8th 1908 – but the main features of the law of the year 1896 remained unaffected. How was insider trading treated under this law? Consider that the term insider did not exist at that time – but contemporary reports of the committee (BEK)[74] that discussed the exchange law pointed out that there was a fear of unregulated speculation that could destabilize the exchange. This fear was caused by the experiences of the crisis in 1873 called "Gruenderkrise" for which a general increase in speculation – among other factors – was made responsible for the pronounced decline in asset prices. What was regarded as speculation? Every transaction motivated by a future increase in the market value of the respective stock was seen as speculation. Thereby, the BEK distinguished between justified and unjustified speculation. A speculation is unjustified if speculators act like gamblers and do not base their decision on the evaluation of the company's economic situation.[75] This definition has obviously nothing to do with our imagination of insider trading. Unjustified speculation is closer to the definition of noise-traders who buy or sell stocks motivated by pseudo-information. They act irrational and are responsible for an additional and not diversifiable stock price risk.[76] Furthermore, the BEK thought that speculators are responsible for the deviation of the current market price of a stock from its justified fundamental value. This excessive over- or underreaction, for instance due to the release of new information, is considered as main source of additional risk and, hence, higher stock price volatility.[77] To reduce the influence of speculation on stock prices, the exchange law prohibited[78] respectively restricted the trade of forward dealings because these financial instruments were seen as a device mostly used by speculators. For instance, future contracts on crops were traded with high volumes at the Berlin stock exchange. To limit the extensive speculation in future contracts on crops was the main aim of the law, whereas the trading of futures with other underlyings like stocks was - under special requirements -

[73] How the exchange law of 1896 affected different lines of business, for instance, the banking industry (see Fohlin, 2002) is still an issue of current research.
[74] See Weber (2000) for a precise discussion of the reports provided by the Börsen Enquete Kommission (BEK).
[75] Even contemporary economists like Bachmann (1898) had problems to precisely distinguish between justified and unjustified speculation.
[76] For a detailed description of noise-trading behavior see Black (1986).
[77] If the exchange law of 1896 really influenced the observable excess volatility is still debating (see, for instance, Wetzel, 1996).
[78] See exchange law (BörsG) §§48-69.

possible.[79] Hence, it was required that a company had to have more than 20 million Mark nominal capital and did not belong to the mining or manufacturing industry. Since the introduction of the exchange law of 1896 and the complete cessation of option dealing in 1914, the futures market regained its former importance as measured by the liquidity and trading volume of the market not before the 1970s.[80] However, insider trading by buying or selling stocks on the spot market was not taken into account.

The disclosure of information as a mean to reduce the information asymmetry on the exchange was improved by imposing requirements that must be fulfilled if the company issued new shares. Hence, the company must reveal publicly all relevant information that could affect the credit risk of the firm.[81] Lacking requirements for the disclosure of other price-sensitive information like a merger announcement is responsible for the observed firm-behavior that some companies delay the declaration of an impending merger. This legal loophole enables to analyze firm-behavior without regulatory restrictions.

3.3.2 Ad-hoc-publication and insider trading in the year 2000

Up to the year 1994, there was a voluntary self-regulation regarding insider-trading. Nowadays, there are several legislative requirements that make insider trading more difficult and costly if trading motivated by an illegal access to private information is detected by the federal financial supervisory authority (BaFin).[82] The Securities Trading Act (WpHG)[83] defines market participants as insiders if they are members of the management or supervisory board respectively large shareholders of the company. Hence, they have access to unpublished information of the company. Every unpublished firm specific information that could have a significant impact on the market value after its public declaration is insider information. Trades based on this information are illegal. However, there is one exception with regard to the work of analysts as long as their analyses are derived from public information. Note that the theoretical model of Kim and Verrecchia (1994) states that one kind of insider trading stems from above-average skills in analyzing public information.

---

[79] Falke (1979) pointed out that the law of the year 1896 was focused mainly on forward dealings regarding grain and flour. He also stressed the new publication requirements if a firm went public or issued additional shares.

[80] In addition, Welcker, Kloy, and Schindler (1992) pointed out that option trading was in turn allowed in 1970. However, between 1st October 1925 and 14th July 1931 option dealing was possible for a short period of time.

[81] See exchange law (BörsG) §36 and §45. However, a violation of these requirements did not trigger remarkable legal consequences – albeit the issuing bank was urged to conduct a thorough proof of the prospectus.

[82] Since 1st May 2002 the BAWe (security supervisory) is embedded into the BaFin.

[83] I refer to the announced version on 9th September 1998. Insider trading is regulated in §12 to §20.

Besides the disallowance of insider trading, the ad-hoc-publication requirement[84] is designed to guarantee that small shareholders (outsiders) can keep pace with better informed market participants. Price-sensitive new information like an impending merger or takeover bid must be announced in a supra-regional stock exchange gazette.[85] The daily newspaper "Handelsblatt" belongs to this group and, hence, I use this source to determine the announcement day.

To discuss the economic impact of a legislative framework on the scale of insider trading, it is worthwhile to assess whether the written laws are really executed. Using the data provided by the BAWe (security supervisory agency),[86] one can argue that the number of investigated insider activities increased between 1995 to 2000 from 24 to 51. After reaching its peak in the year 2001 with 55 cases, the number declined considerably to only 15 cases for the first two quarters of the year 2002. These figures give the impression that insider trading is effectively prohibited – but, in the period 1995 to 2000, only in sixteen cases out of 337 initiated investigations insider activity had legal consequences. Thus, there remain doubts whether the legal framework works as an obstacle to gain extraordinary profits from using insider information.

3.4 Method of Sampling

3.4.1 Determination of the event day

As more thoroughly discussed in chapter two, the day of the announcement of a merger is regarded as event day around which I construct the event window. For the sample of the year 1908, merger announcements published in the daily newspaper `Berliner Boersenzeitung´ between January to June 1908 are considered. The `Berliner Boersenzeitung´ was the leading newspaper for investors; thus, one can regard this released mergers as public information. For the year 1908, I detect 101 mergers of which forty-six are included in my sample because not all companies are listed on the stock exchange. Moreover, thirteen firms decided to hide an impending merger. In contrast to my study, Banerjee and Eckhard (2001) collected 56 mergers of which nineteen firms did not disclose the merger; thereby, the so called first merger wave 1897-1903 is their investigation period. Furthermore, Banerjee and Eckhard

[84] See WpHG §15.
[85] Nowadays, seven daily newspapers fulfill the legal requirements being an official stock exchange gazette. It is also allowed to disclose new information in an electronic system like vwd (see www.vwd.de).
[86] Issues for the year 2000 and the first two quarters of 2002 are available on the homepage (http://www.bafin.de/frame_bawe.htm).

(2001) worked on a weekly level,[87] whereas I use daily returns and determine the exact event-day.

In the year 1908, the involved firms themselves publicly declared impending merger, whereas rumors appeared only seldom in the daily newspaper. One exception was the rumor that an acquirer started buying the stocks of the mining company `Donnersmarksche Huette´. In this case of a hidden merger, it is reasonable to use the day of the publication of the rumor as event day. By reading the newspaper, even uninformed investors can update their information regarding the probability of a tender offer.

For the sample of the year 2000, I use the same methodology and collect 61 mergers by reading the daily newspaper `Handelsblatt´ and by working with online archives.[88] The included mergers were released between January and June 2000. Note that only merger activities in which at least one participating firm is owned by mostly German stock holders are considered. A further requirement is that all stocks have to be listed on German stock exchanges; thereby, the overwhelming part of the included stocks were traded on the Frankfurt stock exchange. In contrast to the historical time period, rumors were very common in the year 2000 and even worse many rumors were false. In this case, false means that the rumor is not followed by concrete negotiations. The sense of these rumors – often spread by the management of the affected firms – is to influence the market value of the firm respectively to confuse competitors. These false rumors are excluded from the sample. Therefore, the event day is the day of the public declaration of an official announcement or a rumor that is followed by negotiations.

3.4.2 Should one include unsuccessful mergers?

In the year 1908, only a few mergers were not successful in the sense that the proposed mergers failed to achieve the necessary majority in the shareholder gathering or legal restrictions prevented the merger. One of the two exceptions was the merger among three shipyard companies `Neptun´, `Howaldswerke´, and `Koch Werft´. In this case, the city of Rostock intervened caused by the fear of loosing an important local tax payer. In addition, the shareholders were not convinced by the `big deal´. To get an impression how this `merger thriller´ was discussed in the daily newspaper, table 3.1 presents the related announcements. The rich details provided by the daily newspaper and the remarkable number of related

[87] They relied on information provided by the weekly newspaper `The Commercial and Financial Chronicle´.

[88] Nowadays, the BaFin provides excellent data on ad-hoc announcements – but this official source neglects rumors that can also affect stock prices tremendously. Hence, one should add the information spread by newspapers to get a detailed impression with regard to the public information.

announcements is comparable to current print media. Henceforth, outsiders in the year 1908 could use the `Berliner Börsenzeitung´ to get up to date information of high quality.

In the year 2000, a remarkable number (22 out of 61) of the announced mergers were not executed later. The merger between `Deutsche Bank´ and `Dresdner Bank´ was among these failures. Why should one include these unaccepted mergers? As shown in chapter two, the dummy variable that takes the value one if the merger gets the approval of the shareholders does not affect the cumulated abnormal return of the acquiring or target firm. Especially in the year 2000, the legal process to accept a merger took several months; hence, the day of the announcement and the confirmation that the merger was in line with the legal framework were timely far away. This means that the market responds to the merger announcement building expectation regarding the acceptance by the monopolies commission, the behavior of the shareholders and so on. Thus, I include declared mergers that are not executed later – but I control for this problem by using a dummy variable or by defining subgroups.

Table 3.1: A historical `merger thriller´ among leading shipyard companies

The chronological announcements indicate the rich flow of newly available information about 100 years ago.

Date of the announcement in the `Berliner Börsenzeitung´	Abbreviated content of the announcement
7th February 1908 Morning issue, insert III, title page	"The three shipyards `Neptun´, `Howaldtswerke´, and `Eiderwerft´ announce a merger. They convene extraordinary shareholder gatherings"
8th February 1908 Evening issue, page 11	"`Howaldtswerke´ call an extraordinary shareholder gathering requesting the approval for the merger with `Neptun´. The gathering will take place on 29th February"
9th February 1908 Sunday issue, insert II	"In a meeting of the advisory board of `Howaldtswerke´, doubts regarding the impending merger arise."
14th February 1908 Evening issue, insert III	"Also the meeting of debenture holders of `Howaldtswerke´ will take place on 29th February"
26th February 1908 Evening issue, page 11	"Strong opposition within the company `Neptun´ forms to argue against the proposed merger. Concerns about disadvantages for the employees of `Neptun´ emerge. In addition, the city of `Rostock´ fears that the headquarters of `Neptun´ could be shifted. Hence, `Rostock´ would loose an important local taxpayer."
29th February 1908 Evening issue, page three	"Shareholder gathering of `Neptun´ rejects the merger for the time being"
1st March 1908 Sunday issue, insert III	"The extraordinary shareholder gathering of `Howaldtswerke´ rejects the merger. This decision is justified by the generally bad shape of the shipyard industry and the expressed disapproval of `Neptun´. Furthermore, the simultaneous merger with more than one shipyard failed. An ordinary shareholder gathering will be conducted on 28th March."
2nd March 1908 Evening issue, page 11	"For the last financial year, `Howaldswerke´ cancels dividend payments"
2nd March 1908 Evening issue, page 12	"General meeting of `Eiderwerft´ rejects the merger with `Howaldswerke´ and `Neptun´ because the `Kochsche Werft´ does not want to participate in the merger."

3.4.3 Determination of the event and estimation period

The event window starts fifteen days before the public announcement and ends fifteen days afterwards. During this predetermined event period, I try to identify the economic impact of a merger by observing the deviation of daily return from the normal stock price movement. Note that the normal return is estimated for the sample of the year 2000 during an estimation period starting at 1st July 1999;[89] thereby, I collect 50 daily returns for each stock and calculate the normal return. Note that the estimation period should be timely sufficiently far away from the event period to avoid that a merger can influence the estimated normal stock price behavior. Any significant deviation from this normal stock price movement serves as a hint that the merger possesses an economic impact on the firms market value. This deviations are called abnormal returns. The choice of the event period for the year 2000 can be justified by calculating the average p-values of the abnormal returns of the whole sample (see figure one). Indicating a high level of stock price movements caused by the merger announcement, the average p-value reaches its minimum around the event day (t=16).[90]

3.5 Event-study analysis

3.5.1 `Recalling´ the basic concept

The aim of event-studies is to detect abnormal returns caused by events like merger announcements. Accordingly, abnormal returns are the deviations of current returns observed during the event period from normal returns. As mentioned above, the normal returns are estimated based on observations collected during an estimation window. Unfortunately, there are six different ways to calculate these normal stock price movements and several additional modifications of these basic concepts are possible.[91]

---

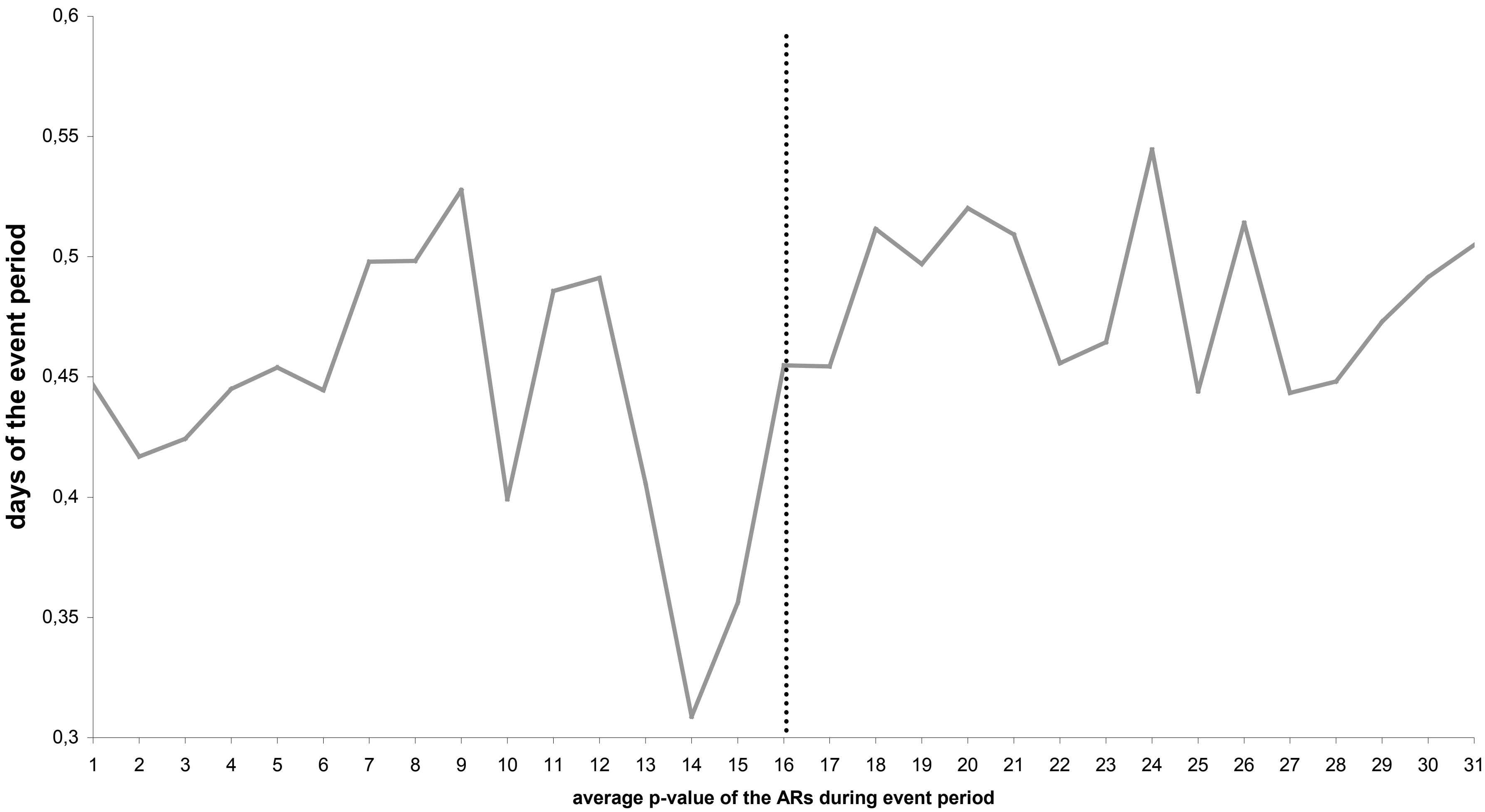
[89] Because many firms in the year 2000 were not listed before July 1999, I decide to start the estimation period later in comparison to my former sample of the year 1908. This problem is of special interest, when mergers among young companies listed on the "Neuer Markt" (a segment for the so called new economy) are taken into consideration.

[90] The picture is similar for the year 1908 as shown in chapter two.

[91] Armitage (1995) provided an excellent overview of the different ways to estimate normal returns.

Figure 3.1: Determining the correct event period for the sample drawn in the year 2000

I plot the average p-value of the abnormal returns (AR) to justify the chosen event period; thereby, the vertical line indicates the event day.



As discussed in chapter two, a simple mean reverting process[92] of returns is appropriate for the historical time period. This assumption leads to the constant mean return model (CMR) and enables to calculate normal returns without using a market index. Because the sample drawn in the year 1908 should be compared to the sample of the year 2000, the calculation of normal returns is based on the CMR model for both samples. Of course, the argument that a trustworthy market index on a daily basis is not available in the year 1908 cannot be used for the later time period. However, if one wants to compare the adaptation process of stock prices due to a merger between these two time periods, it is reasonable to use the same model to determine normal returns. This procedure avoids that the chosen model to estimate normal returns is responsible for detected differences with regard to the abnormal returns around the announcement. Thus, I stick to the simple CMR model and refer to my previous results.[93] Besides the comparability of two different time periods, one should keep in mind that the CMR model and the more sophisticated stochastic market model lead to very similar results as shown in chapter four. Thereby, the portfolio weighted abnormal returns of the whole sample of the year 2000 are obtained using the CMR respectively the market model. Hence, relying on the CMR does not lead to remarkable losses in accuracy of estimating the impact of events.

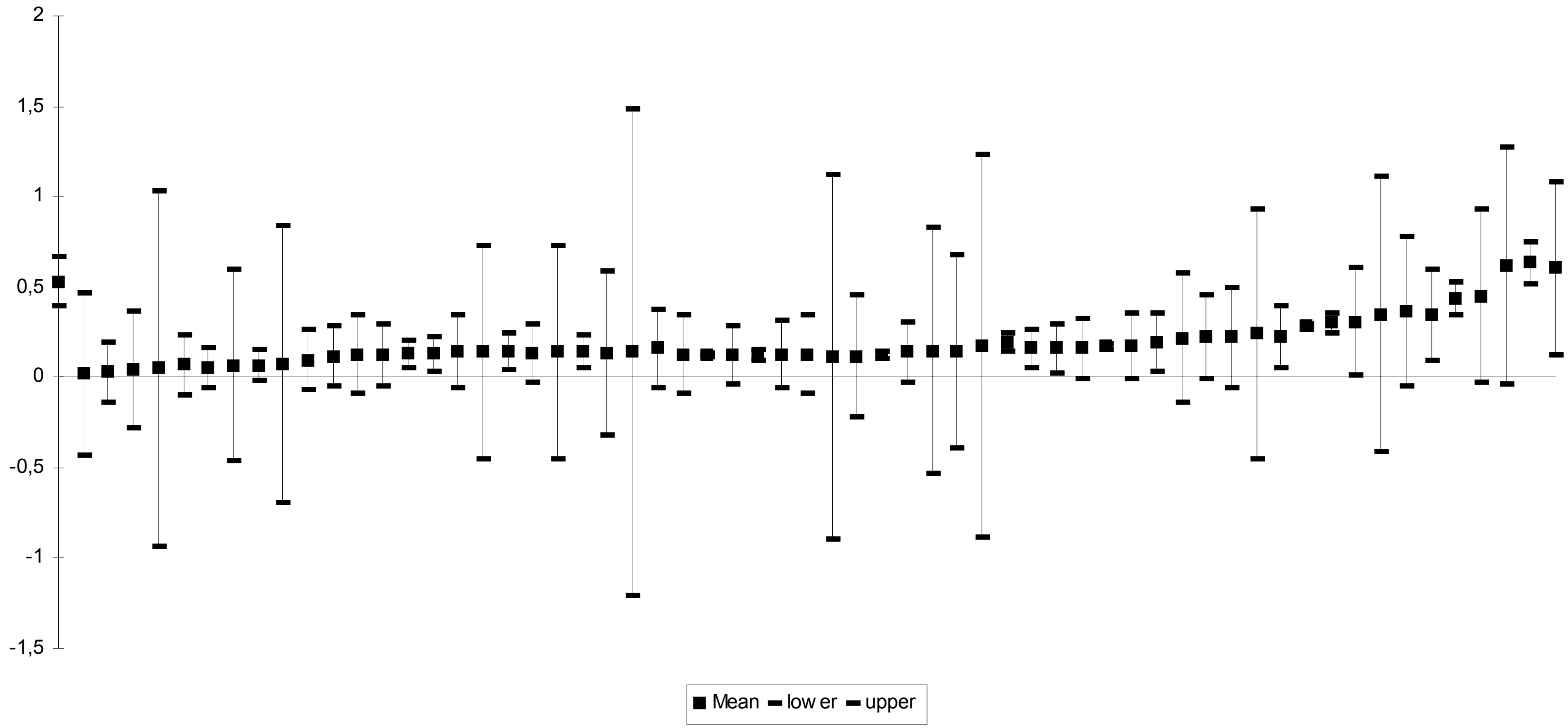
3.5.2 Results of the estimation period in the year 1999

In comparison to the results of the year 1908 (see chapter two), some estimated daily mean returns differ significantly from zero. In general, one can observe a positive drift component of the suggested random walk of daily stock prices. Figure 3.2 plots the 95% confidence interval of the estimated mean returns obtained from the estimation period of the year 1999.

[92] A mean reverting process describes a time series that has a long-term mean.
[93] See chapter two.

Figure 3.2: Estimated mean returns and confidence intervals for the year 1999 of individual stocks

Results from the estimation period contain the upper- and lower-bound of the constructed 95% confidence interval of the estimated mean returns.



3.5.3 Abnormal returns and cumulated abnormal returns in the year 2000

Using the test statistics derived in chapter two as well as the necessary assumptions, I calculate the abnormal returns for each stock and the portfolio weighted average abnormal return for the whole sample and test for significance. Figure 3.3 and table 3.2 show the portfolio weighted abnormal return for each day of the event window and the cumulated portfolio weighted abnormal return; thereby, the time interval over which the daily abnormal returns are added up increases till the whole event period is covered. This cumulated return measures the total change in the market value of the firms triggered by the merger announcement. Figure 3.3 also indicates using gray boxes if the abnormal or cumulated abnormal return is significant on the 10% level of significance. The whole sample consisting of 61 stocks exhibits a decline in stock prices by 3.60% (p-value 0.11) over the whole period of 31 days. However, about three days (t=13) before the announcement day (t=16) the abnormal returns are positive and significantly different from zero. Moreover, distinguishing between executed and prevented mergers enables to assess the knowledge of the market regarding the probability that the declared merger is executed later.

3.5.4 Does the market know if a merger fails to overcome the hurdles?

A large portion of announced mergers that fail to achieve the necessary majority in shareholder gatherings respectively were rejected by the advisory boards could be observed in the year 2000. In addition, restrictive antitrust laws prevented many mergers. In my sample, 22 out of 61 announced mergers were not executed later. The legal framework was completely different to the situation in the year 1908 and, hence, was responsible for the high scale of failed mergers as well as the time intensive process till a merger was accepted by antitrust authorities. Did the market anticipate the failure of mergers? To answer this question, I build up two categories. The first group contains all announced mergers that are later executed, whereas unsuccessful mergers belong to the second group. Then the cumulated

Figure 3.3: Abnormal return and cumulated abnormal return of the whole sample in the year 2000

Figure 3.3 contains the portfolio weighted abnormal returns $\bar{\varepsilon}_t^*$ for each day t∈{1,2,…,31} of the event window and the aggregation over increasing time intervals $\overline{C}(1;\tau_n)$. Gray boxes indicate significance on the 90% confidence level.

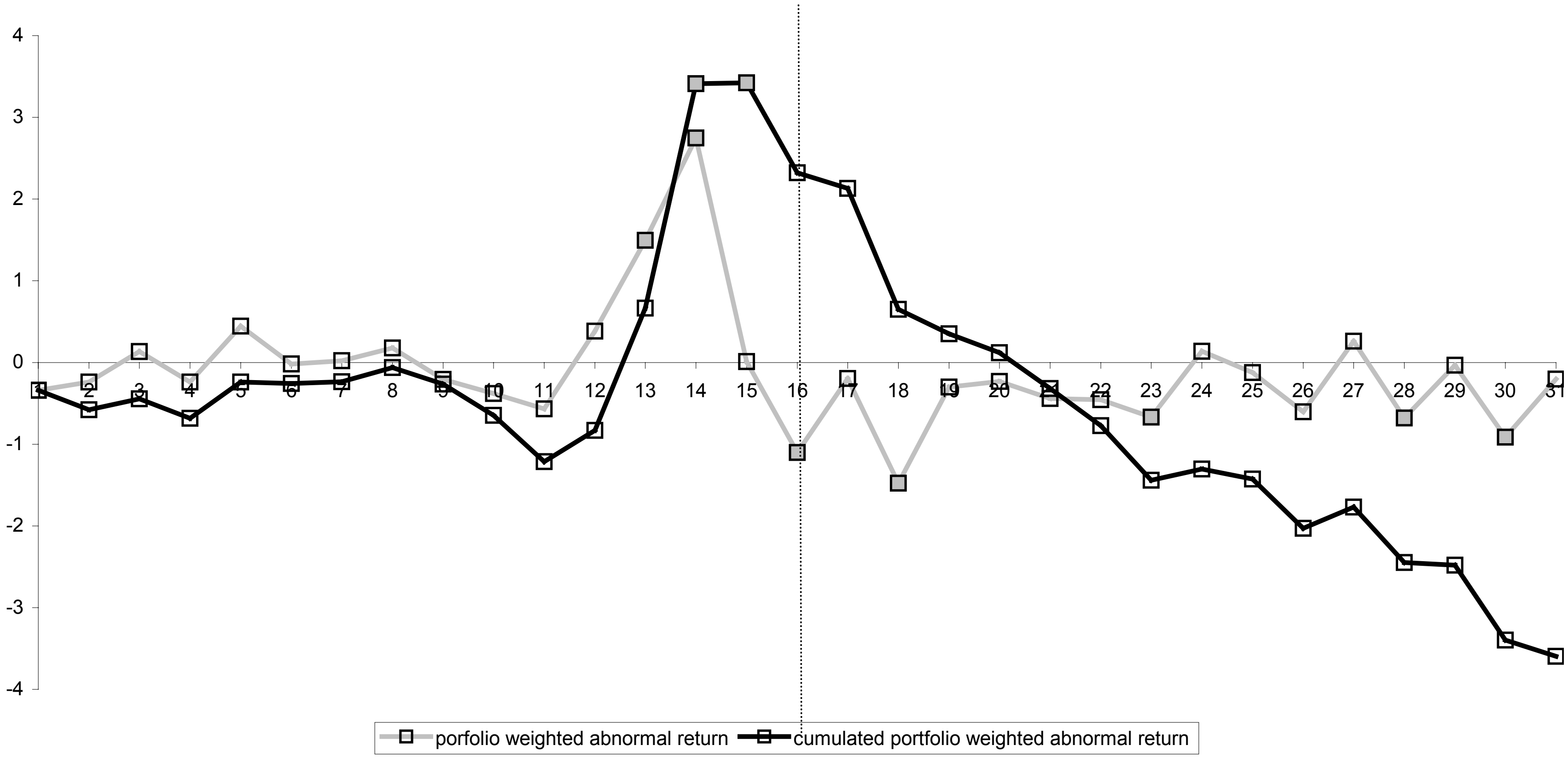


Table 3.2: Abnormal and cumulated abnormal return for the sample of the year 2000

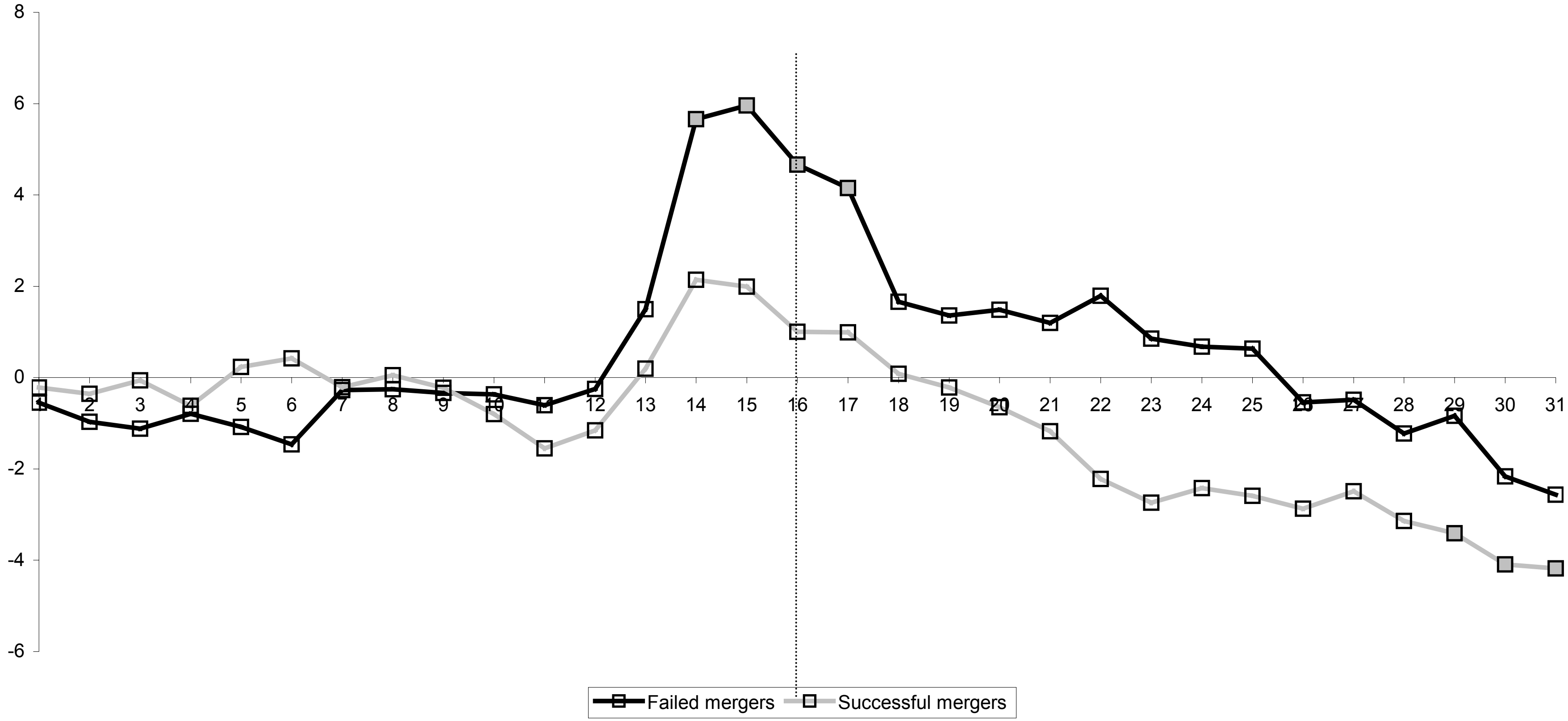
Table 3.2 contains the portfolio weighted abnormal return $\bar{\varepsilon}_t^*$ at each event day t; the third column shows the p-value of $\bar{\varepsilon}_t^*$. The aggregation over different time intervals $\bar{C}(1;\tau_n)$ is listed and the significance is assessed, using p-values that appear in the fifth column. The event day is t=16.

$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\bar{C}(1;\tau_n)$	p-value	$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\bar{C}(1;\tau_n)$	p-value
1	-0.3391	0.402	-0.3391	0.402	17	-0.1940	0.631	2.1288	0.202
2	-0.2369	0.558	-0.5761	0.314	18	-1.4778	0.000	0.6510	0.704
3	0.1329	0.743	-0.4435	0.527	19	-0.2992	0.459	0.3518	0.842
4	-0.2393	0.554	-0.6828	0.399	20	-0.2315	0.567	0.1203	0.947
5	0.4429	0.274	-0.2400	0.791	21	-0.4389	0.278	-0.3187	0.864
6	-0.0178	0.965	-0.2577	0.795	22	-0.4555	0.260	-0.7741	0.683
7	0.0227	0.955	-0.2351	0.826	23	-0.6674	0.099	-1.4415	0.457
8	0.1752	0.665	-0.0599	0.958	24	0.1377	0.733	-1.3038	0.511
9	-0.2040	0.614	-0.2639	0.828	25	-0.1230	0.761	-1.4268	0.481
10	-0.3806	0.347	-0.6445	0.614	26	-0.6022	0.137	-2.0290	0.325
11	-0.5667	0.161	-1.2113	0.367	27	0.2595	0.521	-1.7695	0.400
12	0.3815	0.346	-0.8297	0.554	28	-0.6779	0.094	-2.4473	0.253
13	1.4954	0.000	0.6656	0.648	29	-0.0336	0.934	-2.4809	0.255
14	2.7458	0.000	3.4115	0.024	30	-0.9147	0.024	-3.3957	0.125
15	0.0124	0.975	3.4239	0.029	31	-0.2006	0.620	-3.5963	0.110
16	-1.1011	0.006	2.3228	0.151					

abnormal return is calculated for both subgroups; figure 3.4 shows the results. The adaptation process is very similar between the two subgroups; this empirical finding stresses that the market did not perfectly know whether the merger was later executed. Note that the time span between the declaration of the willingness to merge and the transfer of assets was several months. So to assess if the market reacts to a failed merger, one should use an additional sample constructed around the public announcement of the failure. Note that this declaration is in turn public information. In general, failed mergers were more successful than really executed mergers.

Figure 3.4: Cumulated abnormal return of failed and successful mergers in the year 2000

Figure 3.4 plots the aggregated cumulated abnormal return for increasing intervals starting at t=1 and ranging till t=31; thereby, we divide between mergers that are not executed (failed mergers) and successful ones.



3.5.5 The way of disclosure in the year 1908

In the year 1908, thirteen firms decided to hide information and postponed the public declaration of their willingness to merge, whereas 33 firms revealed their intentions. This firm behavior affects the adaptation pattern of stock prices before and after the event day. Before comparing both groups, the portfolio weighted abnormal return and the cumulated portfolio weighted abnormal returns are calculated. Table 3.3 contains the results for firms that hide information and table 3.4 shows the measures for well-informing firms. Figure 3.5 plots the cumulated effect for both subgroups to evaluate whether the time paths differ.

The stock prices of firms that hide information exhibit a remarkable upsurge over the whole period of 31 days by 5.60% (p-value 0.000), whereas the market values of firms that disclose mergers increase only by 1.63% (p-value 0.027). This in general does not mean that the strategy of hiding influences the success of a merger as measured by the cumulated abnormal return positively. I concentrate on this issue in a subsequent section. It is also likely that important announcements, which possess the capability to change the market value after its declaration tremendously, are hidden by the management of the firms to use this self-created time lag for insider-trading. This impression is confirmed by analyzing the pre-merger gains. These so called run-ups are the cumulated abnormal returns over the pre-event period ($t=1,2,\ldots,15$). Note that $t=16$ is the event day. Following the definition of Banerjee and Eckard (2001) as well as Keown and Pinkerton (1981) that significant price adjustments prior to the public release of a merger are due to insider trading, I compare the pre-event gains between the two ways of disclosure. Table 3.5 contains these run-ups for hidden and disclosed information. If the merger is correctly made public, the cumulated abnormal return prior to the event day adds up to 0.73% (p-value 0.152) and is insignificant. On the event day, the

Table 3.3: Abnormal and cumulated abnormal return for firms that hide mergers

Table 3.3 contains the portfolio weighted abnormal return $\bar{\varepsilon}_t^*$ at each event day t; the third column shows the p-value of $\bar{\varepsilon}_t^*$. The aggregation over different time intervals $\overline{C}(1;\tau_n)$ is listed and the significance is assessed, using p-values that appear in the fifth column. The event day is t=16.

$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\overline{C}(1;\tau_n)$	p-value	$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\overline{C}(1;\tau_n)$	p-value
1	0.1268	0.338	0.1268	0.338	17	0.3929	0.003	5.1725	0.000
2	0.1223	0.355	0.2492	0.183	18	0.1676	0.205	5.3400	0.000
3	0.3979	0.003	0.6470	0.005	19	0.4523	0.001	5.7923	0.000
4	0.1701	0.199	0.8171	0.002	20	-0.0464	0.726	5.7459	0.000
5	0.0616	0.641	0.8787	0.003	21	0.1633	0.217	5.9091	0.000
6	0.2175	0.100	1.0962	0.001	22	0.1818	0.170	6.0909	0.000
7	-0.0423	0.749	1.0539	0.003	23	0.5500	0.000	6.6409	0.000
8	-0.1307	0.323	0.9232	0.014	24	-0.0619	0.640	6.5790	0.000
9	0.1757	0.184	1.0989	0.006	25	-0.0556	0.674	6.5234	0.000
10	0.3001	0.023	1.3990	0.001	26	0.0698	0.598	6.5932	0.000
11	-0.2897	0.029	1.1093	0.011	27	-0.3789	0.004	6.2143	0.000
12	0.8108	0.000	1.9200	0.000	28	-0.2874	0.030	5.9269	0.000
13	2.3959	0.000	4.3159	0.000	29	-0.1946	0.141	5.7322	0.000
14	0.1533	0.246	4.4692	0.000	30	-0.0090	0.946	5.7232	0.000
15	0.3182	0.016	4.7875	0.000	31	-0.1230	0.353	5.6002	0.000
16	-0.0079	0.952	4.7796	0.000					

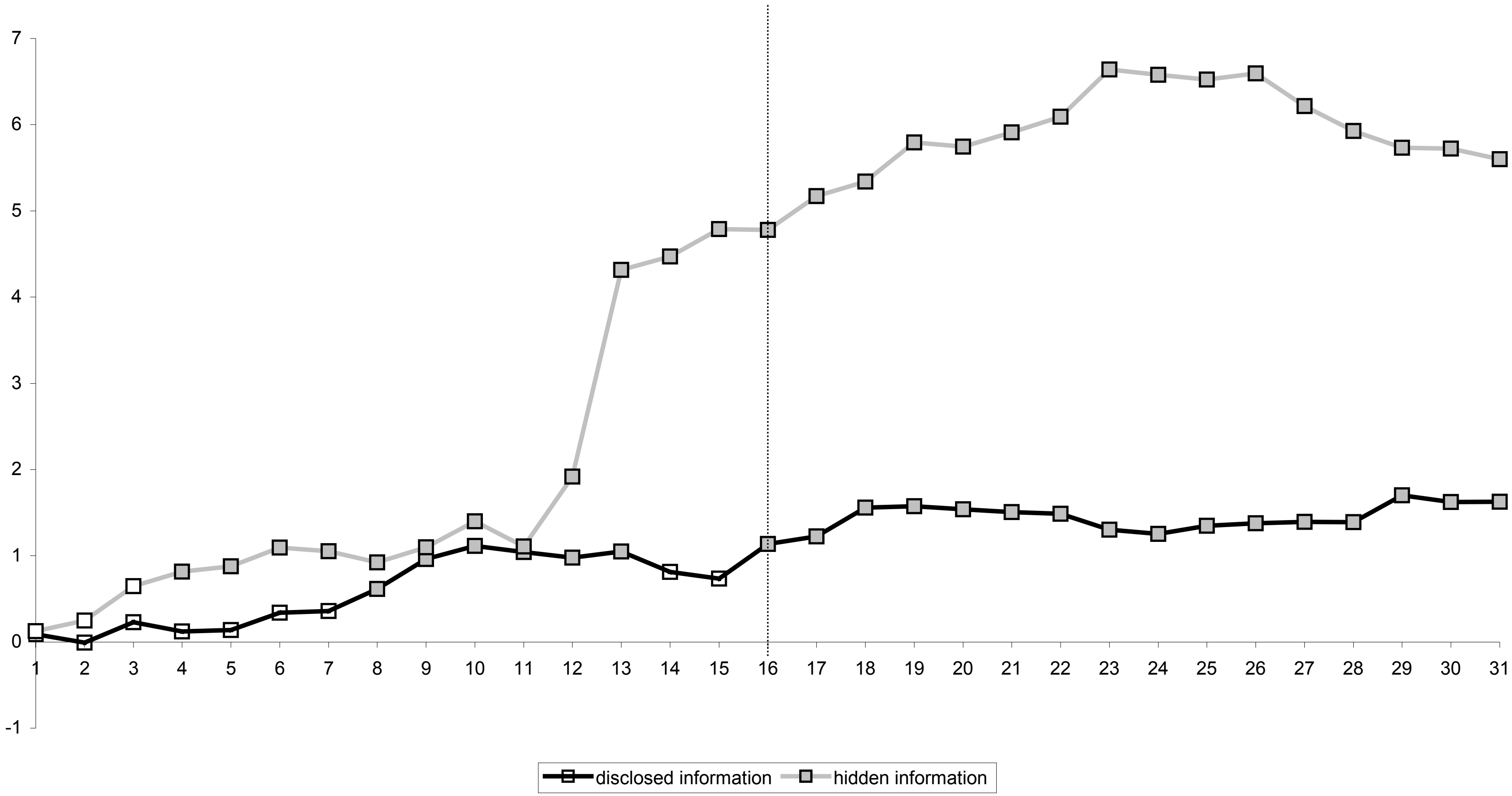
Table 3.4: Abnormal and cumulated abnormal return for firms that disclose mergers

Table 3.4 contains the portfolio weighted abnormal return $\bar{\varepsilon}_t^*$ at each event day t; the third column shows the p-value of $\bar{\varepsilon}_t^*$. The aggregation over different time intervals $\overline{C}(1;\tau_n)$ is listed and the significance is assessed, using p-values that appear in the fifth column. The event day is t=16.

$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\overline{C}(1;\tau_n)$	p-value	$\tau_n = t$	$\bar{\varepsilon}_t^*$	p-value	$\overline{C}(1;\tau_n)$	p-value
1	0.0894	0.499	0.0894	0.499	17	0.0874	0.509	1.2247	0.025
2	-0.0985	0.457	-0.0091	0.961	18	0.3351	0.011	1.5598	0.005
3	0.2384	0.072	0.2293	0.317	19	0.0147	0.912	1.5745	0.006
4	-0.1059	0.424	0.1235	0.641	20	-0.0357	0.787	1.5387	0.009
5	0.0163	0.902	0.1397	0.637	21	-0.0310	0.815	1.5077	0.013
6	0.1989	0.133	0.3386	0.296	22	-0.0192	0.885	1.4885	0.016
7	0.0194	0.883	0.3580	0.306	23	-0.1868	0.158	1.3018	0.040
8	0.2578	0.051	0.6158	0.100	24	-0.0490	0.711	1.2528	0.053
9	0.3476	0.009	0.9634	0.015	25	0.0961	0.468	1.3488	0.041
10	0.1493	0.259	1.1127	0.008	26	0.0295	0.824	1.3783	0.041
11	-0.0704	0.595	1.0424	0.018	27	0.0160	0.904	1.3943	0.043
12	-0.0646	0.626	0.9778	0.033	28	-0.0035	0.979	1.3908	0.047
13	0.0709	0.592	1.0488	0.028	29	0.3089	0.020	1.6998	0.017
14	-0.2359	0.075	0.8128	0.101	30	-0.0748	0.572	1.6249	0.025
15	-0.0793	0.549	0.7335	0.152	31	0.0026	0.984	1.6275	0.027
16	0.4037	0.002	1.1372	0.032					

Figure 3.5: Cumulated abnormal return of firms that disclose respectively hide information in the year 1908

Figure 3.5 plots the aggregated cumulated abnormal return for increasing intervals starting at t=1 and ranging till t=31; thereby, I distinguish between firms that disclose an impending merger and firms that withhold new information from the public.



average market value goes up by 0.40% (p-value 0.032); thus, the announced willingness to merger has a relatively strong impact on the stock prices on the event day. After its revelation, the adaptation process of stock prices is not yet finished, and the cumulated significant effect reaches 0.89% (p-value 0.091). Therefore, it is possible for outsiders by reading the daily newspaper to make profits by buying the stocks of companies involved in merger activities. In contrast, hiding information hurts outsiders because a pronounced upsurge in stock prices by 4.79% (p-value 0.000) occurs during the fifteen days before the merger becomes public information. After outsiders update their information, profits by buying stocks of merging companies shrivel up. Note that the event day has nearly no impact on market values; the abnormal return is very close to zero. Moreover, a considerable part of 85.49% of the whole price impact of a merger is already reflected in the market prices before the release takes place. There is also anticipation of the impending merger if firms do not misbehave – but only 45.07% of the total effect flows into the market prices prior to the announcement.

Table 3.5: Measuring the run-ups for different types of disclosure

Note that p-values appear in parentheses.

	Revelation of information	Hidden information
Pre-announcement gains $t \in \{1,2,\ldots 15\}$	0.7335 (0.152)	4.7875 (0.000)
Gains on the event day $t=16$	0.4037 (0.032)	-0.0079 (0.952)
After-announcement gains $t \in \{16, 17,\ldots 31\}$	0.8940 (0.091)	0.8128 (0.125)
Total change in market value over the 31 days	1.6275 (0.027)	5.6002 (0.000)
Pre-announcement gains in per cent of total change	45.07%	85.49%

3.5.6 Was the regulation of insider trading successful during the last 92 years?

Looking at the figures 3.3, 3.4, and 3.5 gives the impression that nowadays it is only possible making profits by buying in advance of a public announcement. After the event day, the cumulated abnormal returns decline sharply regardless which subgroup is considered. Table 3.6 contains the pre-event respectively after-event changes in market values; thereby, I distinguish among the subgroups: targets, acquiring companies, executed, and prevented mergers. In the two groups of acquiring companies and prevented mergers, remarkable pre-event profits are possible. This gains reach 5.85% (p-value 0.000) in the case of acquiring firms respectively 5.96% (p-value 0.000) if the merger is not undertaken after its declaration. Both aggregated values are highly significant. On the day of the newspaper announcement, acquiring firms loose –1.20% (p-value 0.003) of their market values, whereas prevented mergers show a decrease in stock prices by –1.30% (p-value 0.003). After the event day, both categories exhibit a sharp fall in stock prices by –9.09% (p-value 0.000) and –8.52% (p-value 0.000). So gains from announced mergers are only possible before the newspaper prints the announcement or rumor.

Table 3.6: Pre-event and after-event changes in market values in the year 2000

Note that p-values are set in parentheses.

	Target firms	Acquiring firms	Executed mergers	Prevented mergers
Pre-announcement gains	0.3684	5.8504	1.9928	5.9609
$t \in \{1,2,\ldots 15\}$	(0.813)	(0.000)	(0.179)	(0.000)
Gains on the event day	-0.9796	-1.1976	-0.9894	-1.2991
t=16	(0.015)	(0.003)	(0.010)	(0.003)
After-announcement gains	-4.4135	-9.0902	-6.1721	-8.5237
$t \in \{16,17,\ldots 31\}$	(0.006)	(0.000)	(0.000)	(0.000)
Total change in market	-4.0451	-3.2398	-4.1793	-2.5628
value over 31 days	(0.070)	(0.153)	(0.050)	(0.296)

The two other subgroups, targets and executed mergers, show a different behavior. Pre-merger gains are relatively weak and insignificant; especially, target firms' market values remain nearly unaffected (0.37% with a p-value of 0.813). However, the event day has a strong negative significant impact; hence, targets go down by –0.98% (p-value 0.015) and executed mergers by –0.99% (p-value 0.010). Thereafter, a considerable decline in stock prices takes place. The negative impact of the event day and after-event losses are common features shared by all stocks in the year 2000 – but in some cases pre-event gains are possible. Are these positive cumulated abnormal returns before the official announcement a hint for insider trading as they are in my sample of the year 1908?

If this positive cumulated effect is seen as a result of trading motivated by private information, one has to conclude that 92 years of regulation are worthless in prohibiting insider-trading. If I stick to the argumentation of Banerjee and Eckhard (2001) respectively

Keown and Pinkerton (1981), these uncovered positive run-ups stem from insider activities. But, out of my point of view, there is an obvious difference between the run-ups of the year 1908 and the ones of the year 2000. The run-ups in the historical period correctly anticipate the whole economic impact of the merger, whereas the pre-event adaptation nowadays goes in the opposite direction regarding the total effect of the merger. If one keeps closely to the concept of insider-trading and its influence on the information content of market prices,[94] run-ups that point in the false direction do not convey privately held superior knowledge. Moreover, insider buying stocks of affected companies prior to the public release should make profits in the sense that the market responds during a specific time interval as predicted by insiders. This is not the case in the year 2000 because on the event day all stocks loose on average, despite the positive reaction during the three days before the announcement. Putting this in other words, it states that observing the considerable increase before the event does not help to improve the expectations regarding the whole change in market values triggered by the merger. Therefore, this observed stock price behavior in the year 2000 is due to speculation driven by pseudo-information or irrational trading rules like buy on rumors and sell on facts. Zivney et al. (1996) provided evidence that rumors about impending takeover bids cause a speculative overreaction of the market. Furthermore, Pound and Zeckhauser (1990) found that rumors are followed by strong price reactions – but about half of the published rumors were false in the sense that a merger is not announced later. Note that I excluded false rumors from my sample, maybe taking also false rumors into consideration would provide interesting insights into overreactions of the market. This is a possible extension of my analysis. In addition, I include rumors that are followed by negotiations and a public announcement of a merger. This means that in these cases even before a rumor appears in the newspaper, it spreads and influences the market prices.

3.6 Cross-sectional analysis

3.6.1 Are undisclosed mergers in 1908 more successful?

In this section, I study both potential directions of impact. First, how the way of disclosure influences the total change in market values. Second, how the success of a merger or the expected increase in stock prices affects the decision of a management to uncover information. The results of the previous sections suggest that companies that decide to

[94] Kyle (1985) provided the theoretical foundations of insider-trading and showed that insiders loose parts of their private information by their trades. This loss in their superiority is taken into account when deciding about the order size. This explains the strategic trading behavior of insiders like the splitting of orders, which is also confirmed by empirical studies.

postpone a merger announcement exhibit a stronger upsurge in their market values over the whole event period than do others. To assess whether the requirement to reveal information quickly should be made, I have to take this extraordinary increase in stock prices of undisclosed mergers into consideration. Accordingly, my scope is to figure out whether the way of disclosure influences the success of a merger as measured by the total cumulated abnormal return. If I detected that hiding information yielded extraordinary cumulated abnormal returns, these efficiency gains would be more valuable than the distribution of these efficiency gains between the two groups, namely insiders and outsiders. Note that the detected losses of outsiders are the gains of insiders. These distribution of wealth is without interest if I concentrate on efficiency argumentations – but hurting outsiders can be important for welfare economists. The task is to estimate the partial effect of disclosure on the cumulated abnormal return of a stock during the whole event period by controlling for other stock characteristics. Consequently, I refer to my model specified in chapter two[95] and extend the set of explanatory variables by including a dummy variable that takes the value one if a firm decides to uncover information. The system of equations looks as follows.

$$\hat{C}_i = \beta_0 + \beta_1 \log(cap_i) + \beta_2 Success_i + \beta_3 Change_i + \beta_4 Cash_i + \beta_5 DivGrowth_i + \beta_6 Bank_i + \beta_7 Mining_i + \beta_8 Mean_i + \beta_9 Target_i + \beta_{10} Disclose_i + u_i \tag{3.1}$$

$$Mean_i = \gamma_0 + \gamma_1 \log(cap_i) + \gamma_2 \log(age_i) + \gamma_3 Success_i + \gamma_4 Change_i + \gamma_5 Cash_i + \gamma_6 DivGrowth_i + \gamma_7 Bank_i + \gamma_8 Mining_i + \gamma_9 Target_i + \gamma_{10} Disclose_i + v_i \tag{3.2}$$

Where: β_0… intercept

$Log(cap_i)$… logarithm of the market capitalization as a measure of firm size of firm i

$Log(age_i)$… Age of firm i

$Success_i$… dummy that takes value one if the merger is executed

$Change_i$… dummy that takes value one if the management is replaced

$Cash_i$… dummy that takes value one if the merger is financed by cash.

$DivGrowth_i$… annual growth rate of dividend payments (1906 to 1908)

$Bank_i$… dummy that takes value one if firm i belongs to the banking industry

$Mining_i$… dummy that takes value one if firm i belongs to the mining industry

$Target_i$… dummy that takes value one if firm i is the target of a merger

$Disclose_i$… dummy that takes value one if firm i reveals information

u_i and v_i… are disturbance terms

[95] The discussed endogenity and causality problem that arise using a cross-sectional model and the possible solution through a simultaneous equation model also apply to this paper. Therefore, for a detailed argumentation see chapter two.

The intuition of this model is straightforward. Including the estimated mean, the normal return, into the first equation incorporates the influence of the normal return on the abnormal return. This impact stems from the definition of abnormal returns that are the difference between observed daily returns and normal returns. The estimated mean serves as dependent variable of the second equation; thereby, the second equation is a reduced form. This procedure allows to distinguish between the direct impact of an exogenous explanatory variable on the success of a merger as measured by the cumulated abnormal return and, by using the second equation, the indirect impact of an exogenous explanatory variable on the estimated normal return. Consider that, for instance, firm size has a significant impact on the estimated normal return, and the normal return in turn influences by definition the abnormal return – but firm size has no direct effect on the abnormal return. Applying this system of equations solves this problem and overcomes one pitfall of event-studies. Hence, I avoid the problem that exogenous factors that affect the estimation period are falsely regarded as important factors that explain the success of a merger.

Carrying out this system of equations by using a three-stage-least-squares procedure[96] yields remarkable results that appear in table 3.7. In addition, the set of variables can be reduced further; thereby, the F-test statistic for the imposed restrictions reaches 0.67 with p-value 0.71, and the log-likelihood specification test points in the same direction. Note that the number of firms is 46, thus $i \in \{1, 2, \ldots, 46\}$. To identify this system of equations, I have to exclude one exogenous variable from the first equation.[97] Besides the influential factors – like belonging to the banking industry – that were already detected in chapter two, the dummy variable $disclose_i$ has no direct impact on the success of a merger – but an indirect impact, which is unimportant caused by the lack of causality. Consequently, I conclude that the way of disclosure is not responsible for more profitable mergers. Hence, hiding information does not lead to higher efficiency gains from mergers.

[96] I use a correction of the estimated standard errors for small samples. Furthermore, using system OLS yields to biased estimates caused by the endogenity of the estimated mean; thus, a three-stage-least squares estimation should be applied.

[97] Regardless which exogenous variable is excluded, the results are hardly affected.

Table 3.7: Estimating the impact of disclosure on the total change in market values

Consider that the p-values are set in parentheses.

Three-stage-least squares				
	Equation 1: CARi	Equation 2: $Mean_i$	Equation 1: CARi	Equation 2: $Mean_i$
Constant	4.2274	-0.5564	3.1380	-0.3332
	(0.457)	(0.000)	(0.232)	(0.000)
$Log(cap_i)$	-1.0961	0.0228	-1.1085	0.0137
	(0.152)	(0.139)	(0.121)	(0.386)
$Log(age_i)$	-	0.0997	-	0.0836
		(0.000)		(0.004)
$Success_i$	-2.0515	0.1344	-	-
	(0.632)	(0.172)		
$Change_i$	-0.5771	0.0044	-	-
	(0.841)	(0.948)		
$Cash_i$	1.4990	-0.0593	-	-
	(0.396)	(0.146)		
$DivGrowth_i$	-0.1720	-0.0787	-	-
	(0.946)	(0.168)		
$Bank_i$	4.5502	-0.0629	4.2344	-0.0203
	(0.045)	(0.228)	(0.040)	(0.691)
$Mining_i$	3.5539	-0.1990	3.8883	-0.1515
	(0.341)	(0.003)	(0.271)	(0.023)
$Mean_i$	-53.2898	-	-50.7940	-
	(0.000)		(0.000)	
$Target_i$	0.1928	0.0726	-	-
	(0.925)	(0.123)		
$Disclose_i$	-2.8665	0.1084	-2.7630	0.0468
	(0.235)	(0.032)	(0.158)	(0.311)
"Adjusted R^2"	0.78	0.43	0.77	0.32
F-test (p-value)	7.74	3.46	11.05	4.36
	(0.000)	(0.001)	(0.000)	(0.002)
Observations	46	46	46	46
Specification test:	8.58			
Log-likelihood ratio	(0.572)			

But when one regresses the same set of explanatory variables only on the pre-announcement gains as captured in the cumulated abnormal return up to time t=15, the impression changes. Table 3.8 presents the outcomes for the run-ups. Obviously, the way of disclosure affects the pre-announcement stock price movement. Within the next section, I discuss the factors that could motivate the manager to hide information.

Table 3.8: Estimating the impact of disclosure on the pre-announcement gains

Consider that the p-values are set in parentheses.

Three-stage-least squares	Equation 1: CARi	Equation 2: $Mean_i$	Equation 1: CARi	Equation 2: $Mean_i$
Constant	4.0801	-0.5564	4.2014	-0.4574
	(0.196)	(0.000)	(0.102)	(0.002)
Log(cap_i)	-0.0432	0.0228	-	-
	(0.918)	(0.139)		
Log(age_i)	-	0.0997	-	0.0983
		(0.000)		(0.001)
$Success_i$	-4.6132	0.1344	-4.5643	0.0931
	(0.055)	(0.172)	(0.047)	(0.368)
$Change_i$	-0.2701	0.0044	-	-
	(0.865)	(0.948)		
$Cash_i$	0.3231	-0.0593	-	-
	(0.741)	(0.146)		
$DivGrowth_i$	-0.5296	-0.0787	-	-
	(0.705)	(0.168)		
$Bank_i$	2.1994	-0.0629	2.0806	-0.0196
	(0.079)	(0.228)	(0.043)	(0.682)
$Mining_i$	1.6917	-0.1990	1.7288	-0.1558
	(0.412)	(0.003)	(0.300)	(0.018)
$Mean_i$	-42.7989	-	-42.7364	-
	(0.000)		(0.000)	
$Target_i$	2.1881	0.0726	2.1516	0.0458
	(0.058)	(0.123)	(0.036)	(0.337)
$Disclose_i$	-2.1111	0.1084	-2.1699	0.0662
	(0.115)	(0.032)	(0.053)	(0.190)
"Adjusted R^2"	0.86	0.43	0.86	0.33
F-test (p-value)	15.08	3.46	20.60	3.80
	(0.000)	(0.001)	(0.000)	(0.002)
Observations	46	46	46	46
Specification test:	7.65			
Log-likelihood ratio	(0.468)			

3.6.2 If managers expect profitable mergers, they will hide information

I should stress that it is unknown at which point in time the management of a company decides to overtake a competitor as well as the scale of the time lag between the decision to merge and the public announcement of a merger. Thus, my logit models that try to explain at which point in time the decision to disclose or hide ought to be made is only an attempt to detect unobserved decision processes within a company. However, this stylized model helps to shed some light on this issue – but it should not be over-interpreted. The influence of the cumulated abnormal return respectively of its expected value on the way of disclosure is now at the core of my analysis. Now, the disclosure is the dependent variable, whereas the cumulated abnormal return serves, besides the above used set of variables, as additional influential factor.

$$Disclose_i = \beta_0 + \beta_1 \log(cap_i) + \beta_2 Success_i + \beta_3 Change_i + \beta_4 Cash_i + \beta_5 DivGrowth_i + \beta_6 Bank_i + \beta_7 Mining_i + \beta_8 Mean_i + \beta_9 \hat{C}_i + u_i \tag{3.3}$$

Before estimating this equation applying a logit regression, one should specify the degree of knowledge regarding the total cumulated abnormal return. Does the manager of a company know how large the whole economic impact of a merger is? Taking the results presented in table 3.5, one can argue that a large part (85.49%) of the total cumulated effect is already anticipated before the merger is made public. Thus, it seems to be plausible to assume that a manager can anticipate a relatively large part of the change in market. To consider different levels of the manager's knowledge, I run regression (3.3) with different specifications of the cumulated abnormal return. Hence, table 3.9 contains five alternative outcomes of equation (3.3); thereby, the degree of knowledge is reduced starting from knowing the cumulated abnormal return for sure in t=31 to observing the cumulated abnormal return four days before the announcement t=12. Note that the dummies $change_i$ and $success_i$ are dropped from equation (3.3) because all executed mergers after which the management was replaced disclosed their willingness to merge. Thus, including these variable enables one to make a

Table 3.9: What influences the decision to uncover the willingness to merge?

Robust p-values appear in parenthesis. The estimates are coefficients not odds ratios. CAR is the abbreviation for cumulated abnormal return. To calculate the number of correctly classified firms, I use 0.5 as cutoff rate.

	Manager anticipates CAR in t=31 perfectly		Manager anticipates CAR in t=15 perfectly		Manager anticipates CAR in t=14 perfectly		Manager observes CAR in t=13 perfectly		Manager observes CAR in t=12 perfectly	
Constant	4.9313	(0.025)	5.4373	(0.013)	5.1220	(0.017)	4.9390	(0.026)	3.3051	(0.097)
CAR for different t	-0.0805	(0.010)	-0.1325	(0.013)	-0.1130	(0.021)	-0.1028	(0.052)	-0.0663	(0.460)
Log(cap)	-0.7302	(0.015)	-0.6716	(0.018)	-0.6424	(0.019)	-0.6428	(0.019)	-0.5626	(0.038)
Log(age)	-0.7394	(0.196)	-0.9568	(0.109)	-0.8766	(0.135)	-0.8001	(0.188)	-0.3056	(0.592)
Cash	0.4263	(0.664)	0.3374	(0.741)	0.2846	(0.774)	0.2193	(0.822)	0.0016	(0.999)
Target	-2.8119	(0.013)	-2.6004	(0.027)	-2.5769	(0.024)	-2.5898	(0.026)	-2.4971	(0.019)
DivGrowth	2.4850	(0.131)	2.5771	(0.124)	2.5488	(0.128)	2.5220	(0.132)	2.1004	(0.158)
Bank	3.0104	(0.011)	2.9107	(0.022)	2.8882	(0.017)	2.8260	(0.020)	2.5640	(0.018)
Mining	3.0252	(0.059)	3.2524	(0.092)	2.8761	(0.101)	2.7258	(0.110)	1.9037	(0.125)
Pseudo R^2	0.34		0.35		0.34		0.32		0.28	
Wald Chi^2 (p-value)	15.50	(0.050)	13.64	(0.092)	12.88	(0.116)	11.30	(0.186)	10.83	(0.211)
Observations	46		46		46		46		46	
Per cent of correctly classified firms	82.16%		82.61%		82.61%		82.61%		80.43%	

perfect prediction about the possibility of disclosure. This would prevent the likelihood optimization technique to uncover the influence of other explanatory variables; thus, these two dummies are excluded.

Of course, it is impossible to state that the manager knows with confidence the total effect of a merger announcement on the market value. So the results for the information set at t=31 are based on the assumption of a perfect anticipation. Besides this problem, the outcomes underlines that the higher the cumulated abnormal return in t=31 the smaller the probability that the merger is disclosed. Despite my doubts about the highest level of anticipatory power, I stress that the event-day is not determined by the management. In contrast, a published rumor forces the management to nail their colors to the mast. So it seems to be possible that managers trade shares on the open market using their superior information that they will not disclose the merger. To avoid that their insider-trading yields to too pronounced price reactions, they act cautiously and split up their desired trading volumes into small pieces.[98] Now a rumor about their trading activities is published and makes their superior knowledge worthless. Thus, it seems to be likely that managers anticipate higher cumulated abnormal returns than the pre-event gains might suggest because their cautious trading prevents the market price to reflect the full degree of their private information.

Even if one rejects this assumption and refers to the information set at t=15 and carries out the regression (3.3) using only the pre-event gains, the results stay nearly unchanged. If stock prices exhibit a strong increase during the last fifteen days, the probability of disclosure is diminished. In addition, the larger the company as measured by the market capitalization the smaller the desire to uncover information. Companies that are targets of a takeover have in general weaker incentives to disclose, whereas mining and banking companies reveal information. This negative impact on the possibility to disclose of the cumulated abnormal return is observable till three days before the event day (t=13). For days that are further away the effect of the cumulated abnormal return disappears. Moreover, the explanatory power indicated by the pseudo R^2 and the Chi^2 statistic is reduced the lower the assumed level of knowledge.

Note that anticipating the cumulated abnormal return three days before the public declaration has not to be a real anticipation. If the decision to disclose respectively hide the merger is made only three days before it becomes public, managers can observe the cumulated effect till t=13. Based on their observations, they can decide to postpone the announcement to use their information to earn extraordinary profits.

[98] The splitting of trading volumes by insiders follows the logic of the Kyle model (1985) and can be observed in empirical research (see, for instance, Chan and Lakonishok (1995)).

My hypothesis that the way of disclosure is closely linked to the way of financing a merger is not confirmed by these results (see table 3.9) regardless which level of knowledge one suggests. The dummy variable $cash_i$ is always insignificant and the coefficients are positive.

3.7 Conclusion

I uncover that hiding information hurts outsiders and is responsible for the strong anticipation of the whole economic impact of a merger. This is reflected in the current market prices on the event day. I attribute the detected run-ups in the year 1908 to insider-trading because this pre-event adaptation of stock prices points in the correct direction in comparison to the whole economic impact of a merger. Thus, a large part of the change in market values is already anticipated prior to the public release. This means that by the trading of better informed market participants new information about the true underlying fundamental value of the company, after announcing the merger, is reflected in the current stock prices. Thus, the order stream of insiders conveys information. This is in line with the theoretical model constructed by Kyle (1985).

In contrast, the run-ups in the year 2000 point in the false direction regardless which group of stocks is considered. Pre-event gains are followed by pronounced declines in stock prices starting on the event-day. Because the stock prices prior to the event-day do not reflect private information regarding the future losses, I consider this pre-event movements as speculative over-reaction. This stock price behavior may be driven by trading rules like buy on rumors and sell on facts.

Using cross-sectional models, the direction of influence between the cumulated abnormal return and the way of disclosure is analyzed. To reveal an impending merger does not lead to higher cumulated abnormal returns, when one controls for other stock characteristics. Consequently, hiding information yields no additional efficiency gains from mergers.

Therefore, I can concentrate on the distributive effect of hiding mergers on insider and outsider gains. This finding enables me to make a clear statement that forcing companies to uncover new information protects outsiders and does not affect the efficiency of takeovers.

Depending on the level of knowledge, the cumulated abnormal return has a negative impact on the probability that a merger is made public in advance. The stronger the expected market response triggered by a merger the larger the incentives to hide information. In addition, a mechanism of self-regulation is not confirmed by my data. Voluntary disclosure that facilitates to raise up money for the expansion of a firm cannot be observed. Hence, intervention of the state is necessary. Obviously, dividing between cash payment and financing a merger by issuing new shares is very crude. Thus, using more precise variables that indicate the dependence on outsiders, as financiers, seems to be worthwhile for further research.

4. The limitation of event-study analysis: Problems and alternative methods

4.1 Extended abstract

This chapter seeks for answers to the following questions: Does the choice of the estimation period matter? How important are exogenous and time dependent shocks? Is it possible to detect abnormal returns in situations without considerable price-sensitive events? Does the non-synchronous trading in historical time periods affect the results? Using a panel study and CUSUM (cumulated sum of residuals) tests to detect structural breaks, I try to explain the process of abnormal return – but an intervention model seems to be more appropriate. Furthermore, modeling the volatility of share prices using a GARCH approach uncovers the event induced uncertainty around announced mergers.

4.2 Introduction

Due to the variety of issues I cover in this chapter, it is hardly possible to write a single introduction. Henceforth, I concentrate on summarizing the main questions of this chapter. A broad literature emerged[99] that highlighted the problems inherent with event-studies; thus, I point out some specific issues that are most likely to cause biases. Accordingly, this paper serves as a litmus-test whether my former results are stable if essential assumptions are violated. Of course, it is necessary to impose these assumptions to derive the test statistics that detect abnormal returns, and these assumptions are widely accepted in the literature. Nevertheless, I tackle these problems re-considering my own results in a critical manner. Such a discussion can also be a breeding-ground for new ideas that help to overcome observed traps of event-study methods. I especially focus on the following points: Does the choice of the estimation period matter? How important are exogenous and time dependent shocks? Is it possible to detect abnormal returns in situations without considerable price-sensitive events? In addition, I provide evidence whether the choice of the model to estimate normal returns matter. Besides this merely technical issues, one should have in mind that focusing on the increase of the cumulated abnormal return is nothing else but shareholder value maximization. Jarrell et al. (1988) reviewed some alternative concepts which explicitly take into account the needs of other stakeholders.

Thus far, I have not precisely modeled the whole adaptation process of stock prices. Hence, I put some efforts into this task; however, it turns out that traditional panel based regression analysis fails to achieve the goal. Correspondingly, I modify the transfer function

[99] Armitage (1995) provided an excellent review paper.

approach and try to distinguish between stochastic fluctuations in daily returns and the deterministic part due to the merger announcement. Despite confirming my former results, the transfer function approach itself possesses several weak points.

Event-induced uncertainty and with that time-varying conditional variance is a well known phenomenon of nowadays stock markets. Can I also apply ARCH respectively GARCH models to increase my understanding regarding the influence of mergers on the volatility in historical periods? Based on my results and Savickas's (2003) corrected test statistics for event-studies, I can incorporate increases in volatility due to announcements into an additional term for standardizing abnormal returns. This can correct my derived test statistics for the event-study.

Working with historical data, thin markets and corresponding liquidity risks should not be neglected. Furthermore, using daily closing prices usually yields a non-continuous trading pattern, which can affect some crucial time series properties of observed daily returns. Trading pauses could bias my test statistic and my former results – but based on information provided by daily newspapers, it is possible to assess the importance of this problem.

Thereafter, I conclude and discuss the general weaknesses of the short-term analysis that I conducted thus far. Motivated by these detected shortcomings, the fifth chapter introduces the long-term analysis of mergers.

4.3 General problems using the event-study approach

4.3.1 How important is the length L of the estimation window?

In my study on the merger paradox in the year 1908, I used fifty observations of daily returns starting on 1st January 1907 of every stock in my sample. To assess the impact of the length of the event window, the number of observations is now doubled. The variance of the estimated mean $Var(\hat{\mu})$ should be reduced by approximately one half if the assumptions regarding normally identically independently distributed daily returns is fulfilled to a high extent. Note that this variance term is in general less important than the variance of the error terms of the constant-mean-return model (CMR) σ_e^2 (see 2.10). After doubling the length of the estimation period, the variance of the estimated normal return $Var(\hat{\mu})$ increased on average by 16.29% – although one should observe the opposite reaction. Accordingly, the i.i.d. assumption imposed on daily returns is far from being true. In addition, the total average increase in the additive variance term is 95.33% because the variance of the error term of the CMR model[100] skyrocket caused by stronger deviations of returns from their estimated mean. Therefore, I

[100] See equation (2.4).

conclude that increasing the estimation window does not lead to an improved estimation of the normal return because the range of the normal stock price movement is even larger. Henceforth, it would be more difficult to detect abnormal share price movements.[101]

Besides different variances, estimated normal returns may differ considerably; thus, abnormal returns could be severely influenced by the chosen estimation period. For instance, Klein and Rosenfeld (1987) argued that bull or bear markets can bias the estimation of normal returns. Accordingly, I should investigate whether changes regarding the estimation period possess a remarkable impact on normal returns. Sticking to the standard procedure discussed by Levin (1999),[102] I test whether the normal returns based on the period from January to February 1907 differ significantly from the enlarged period that starts in January ending after 100 daily observations. The t-values are generally very small – but in one case out of forty-five the t-value reaches –3.31; however, a considerable distortion of my former results can be ruled out. Note that the exceptional case of `Magdeburger Privatbank´ also shows that using the period from January to February 1907 yields significantly higher normal returns by 0.09 percentage points than choosing the extended period. Consequently, extending the period would in this case lead to higher abnormal returns.

One can also assess whether the mean return deviates for the whole sample of companies if the event period is changed. Consequently, the estimation period from January to February 1907 is replaced by the following two months, March and April 1907. Hotelling's T-squared reaches 68.47 and the corresponding F-statistic 0.84 (p-value: 0.727); hence, the null hypothesis that mean returns do not differ for all companies cannot be rejected. Based on these results, the choice of the estimation period and its length seem to be of minor importance for my former outcomes.

4.3.2 Is it possible to detect abnormal returns in time periods without events?

Using the estimation period, January to February 1907, with fifty observed daily returns for every stock, I test for abnormal returns by defining the following two month with fifty daily returns as event period. Note that during this defined event period the included firms did not announce any mergers;[103] thus, one should expect that abnormal returns cannot be detected. In

[101] Note that this empirical finding is sample specific and can be explained by the strong decline in share prices from March 1907 to the end of the year 1907.

[102] Levin (1999) is an excellent introduction into these test procedures. I use a standard test statistic that allows different variances in both estimation periods; thereby, degrees of freedom for the test are obtained from Welch's approximation formula. Nevertheless, Satterthwaite's formula yield similar outcomes.

[103] One can check this easily by reading the annual information provided by the `Handbuch der deutschen Aktiengesellschaften´; however, very small transactions, e.g. acquisition of another company's branch in a specific city, are not always reported.

contrast to my former event studies, I use the same event period with respect to the calendar time for every stock in my sample. This procedure uncovers one essential weakness of the event study approach; event studies do not belong to time series methods because event period and calendar time differ.

Furthermore, event studies work best if the event periods of the respective stocks are seldom overlapping. The reason for this finding is straightforward. An event study that is based on event periods that are equal to a specific calendar period cannot distinguish between abnormal returns triggered by remarkable stock specific events and erratic time shocks that usually hit all stock at the same time.[104] Because the inference is usually based on portfolio weighted abnormal returns, a exogenous time shock that affects all stocks in the market at the same point in time causes significant abnormal return.

Overlapping event periods, hence, lead to a so called clustering (see Armitage, 1995) of cross-sectional units. Brown and Warner (1980, 1985) provided empirical evidence of the importance of this problem using Monte Carlo studies. Clustering of observations also leads to cross-correlation of abnormal returns; therefore, critical assumptions imposed to derive my test statistics in chapter two are violated.

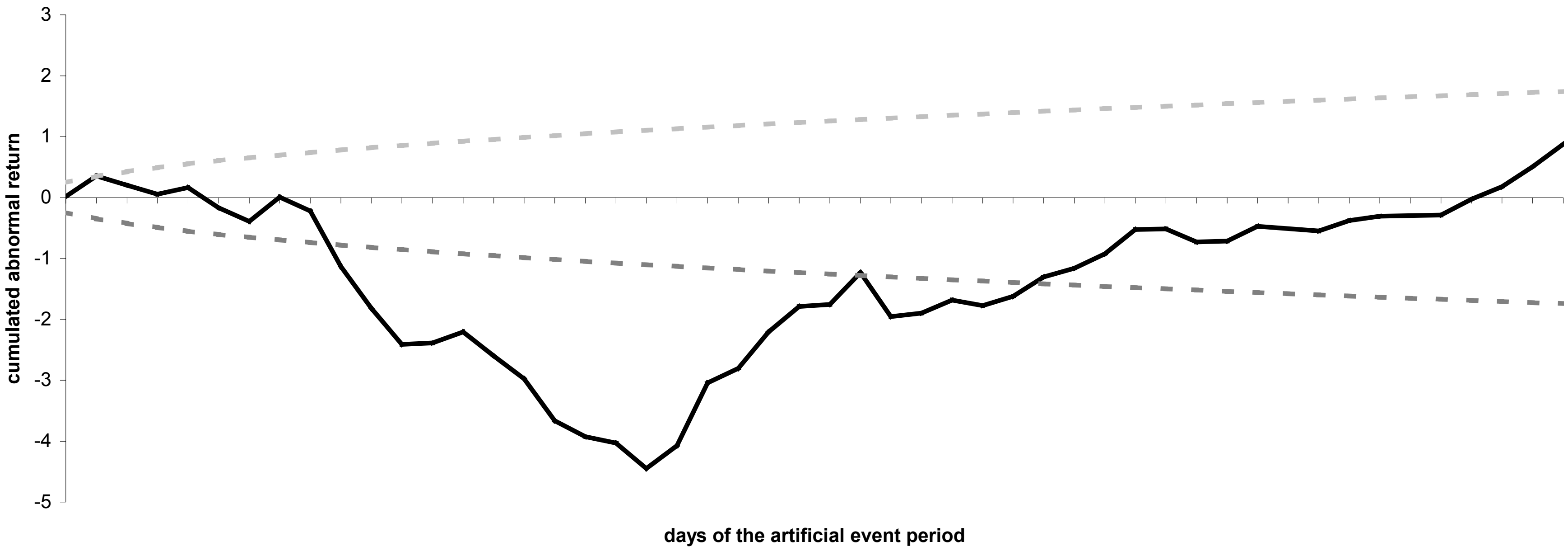
Figure 4.1 plots the cumulated portfolio weighted abnormal return of my artificial event period. Even without price-sensitive events on the micro-level, exogenous time shocks caused a significant deviation of current daily returns from their predicted normal stock price movement. In addition, this deviation is only a temporary perturbation.

---

[104] For instance, unexpected macroeconomic shocks. Chapter five discusses this point thoroughly.

Figure 4.1: Time shocks as artificial events

Note that the artificial event period covers 50 days. Besides the portfolio weighted cumulated abnormal return, the upper and lower critical values are plotted. If the cumulated fluctuation process exceeds one of this boundaries, the deviation is recognized as being significant on the 95% level of significance.



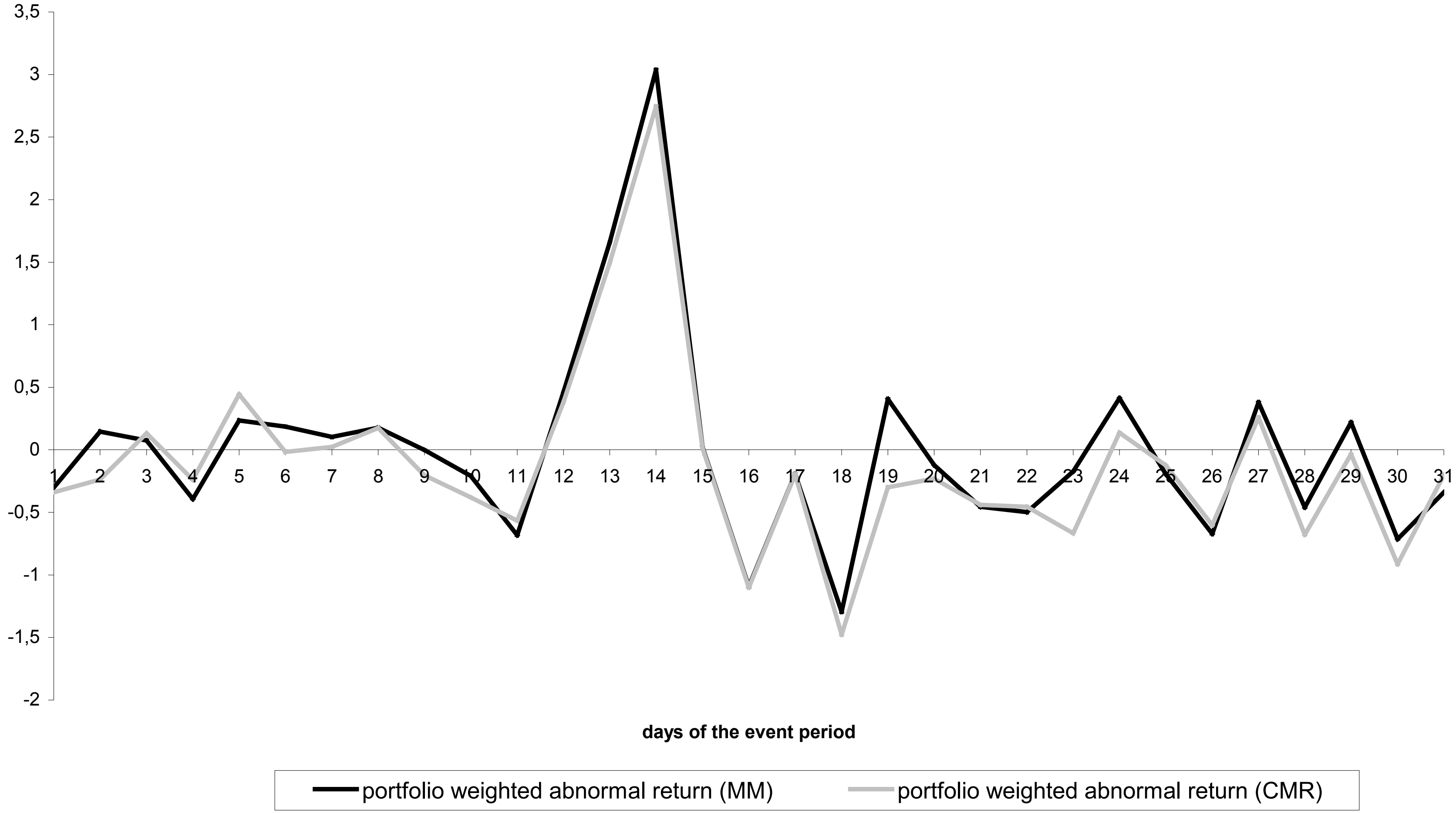
4.3.3 Is the constant-mean-return model inferior in comparison to the market model

As mentioned in former chapters, there are indeed good reasons to choose the constant-mean-return model (CMR) to estimate the normal return in historical periods. But for the sample of the year 2000, a market index is easily available; hence, one can prove whether the market model (MM) leads to better estimates than the simple CMR model. Hence, I carry out the stochastic market model[105] that is widely accepted and used nowadays. The daily return of stock i at time t denoted R_{it} serves as dependent variable, whereas the daily return of the market index DAX30 R_{mt} enters the equation as explanatory variable. A stock specific constant term is also included. After deriving the normal return, the deviation of current daily returns from its predicted level measures the abnormal performance. To compare the CMR and the MM model, figure 4.2 plots for both versions portfolio weighted abnormal returns. The differences between these two sequences are negligible; thus, all results obtained in my former studies stay valid if the market model determines the normal stock price movement. This finding is in line with the simulation experiments of Brown and Warner (1980, 1985).

[105] This model version does not impose any further restrictions on the parameters like theoretical market models do (see Dimson and Marsh, 1984). Since Fama et al. (1969) who applied the simple stochastic market model for the first time, it became the most common way to determine normal returns.

Figure 4.2: The `performance´ of the constant-mean-return (CMR) and the stochastic market model (MM)

To evaluate which model outperforms the other, portfolio weighted abnormal returns are calculated based on the CMR or the MM; thereby, I use the sample drawn in 2000.



4.3.4 Do dividend payments affect the results?

During the event period, only one company, namely the mining firm `Laurahütte´, issued a quarterly dividend payment; hence, stock prices on the fifteenth day after the announcement are quoted as ex-dividend prices. In my former event-studies presented in chapter two and three, I worked with these ex dividend stock prices because the dividend payment seems to be negligible. Note that the stock price reached 215 Mark, whereas only a quarter of the annual dividend of 12% with respect to the nominal capital N_{it} was issued. One method to incorporate dividend payments is to add them to the ex-dividend stock price. This requires to know the day of the issue of dividends. If the date is uncertain, an alternative approach corrects daily returns with regard to annual dividend payments D_{it}; thereby, one calculates the theoretical daily dividend that should be reflected in current market prices $P_{it.}$[106]

$$R_{it}^{Div} = R_{it} \cdot \frac{P_{it}}{P_{it-1}} \left(1 + \frac{D_{it} \cdot N_{it}}{100} \cdot \frac{1}{260} \right) \qquad (4.1)$$

The difference between the observed daily return R_{it} and the daily return with embedded dividends R_{it}^{Div} is relatively low. During the event period the maximum deviation of both time series reaches 0.0255%. Obviously, the impact of dividends becomes more essential if one aggregates daily returns over increasing time intervals. Deriving the cumulated abnormal return for the whole event period with and without controlling for dividends uncovers an aggregated underestimation by neglecting dividends of about 0.1211%. In comparison to the total change of the market value of 2.42%, this underestimation seems acceptable. Besides the distortion during the event period, the estimated normal return is also understated if dividends are not taken into account. In the case of `Laurahütte´, the estimated mean return with adjustments for dividends is 0.0025% higher than without dividends. Hence, the normal return should be 0.0025% higher which diminishes the cumulated abnormal return over the whole event period by 0.0775%. Note that this reduction of cumulated returns caused by higher normal returns is outweighed by the underestimation of cumulated effects of about 0.1211% if dividends are neglected. Therefore, controlling for dividend payments yields to higher cumulated abnormal returns of about 0.0436%. From my point of view, this difference is unimportant and does not justify to spend much time on controlling for dividends. Furthermore, this example underlines that the rejection of the merger paradox is even more likely if dividends are considered.

[106] A similar formula was applied to US and Canadian securities by `Datastream´ until 1973 because detailed information on the exact dates of dividend payments was not accumulated. Accordingly, this method works very well with historical data if one knows only annual dividend payments. Note that I use 260 trading days for this calculation. However, the number of days is not essential for my statements.

Besides this merely technical issue, one should also have in mind the institutional situation during the pre-World-War I period. According to the exchange law of the year 1896,[107] share prices should not reflect `interest rates´ for holding an asset because every shareholder received a 4% annual interest rate payment over the holding period. After the company issued dividends, this prior interest payment are subtracted from dividends. Such pre-payments make a correction for dividends even more complicated and, as discussed above, potential distortions are of minor importance.

4.3.5 Shareholder value orientation versus redistribution theory

Jarrell et al. (1988) highlighted that detecting an increase in market values of acquiring and target firms is certainly not enough to call a merger successful. Obviously, my approach defines a successful transaction as one that maximizes the shareholder value of the involved companies. Only a few empirical studies focused on other stakeholders that could be influenced by mergers, namely employees and bondholders. Nevertheless, the results are ambiguous,[108] and due to a lack of information, I cannot carry out similar studies for the pre-World-War I period.

4.4 Explaining the adaptation process of stock prices: A traditional view

4.4.1 Extension of the cross-sectional model

The following sections discuss an extension of my former cross-sectional models presented in chapter two and three. Restricting myself to explaining only cumulated effects of mergers does not exploit my full knowledge regarding the adaptation process of share prices. Accordingly, by applying a simple panel regression, I try to explain the whole adaptation process of each stock during the event period using stock characteristics.

Inherent with a panel based approach is the problem of clustered observations; thereby, clustering means that observations within a cluster are very similar, whereas among clusters observations are independent. Stocks exhibit time paths of abnormal returns that possess a similar shape among each other. Furthermore, my panel data set contains explanatory variables like the line of business that are constant over the 31 days of the event window. Therefore, residuals of such a panel regression follow similar patterns like abnormal

[107] This regulation came into force at the beginning of the year 1899. Note that, for instance, the stock prices in Paris and London reflected an implicit interest payment for holding the asset.

[108] Denis and McConnell (1986) could not find any evidence that bondholders are affected by mergers even if the transaction is financed by cash payments. Shleifer, Summers (1987), Brown and Medoff (1987) provided ambiguous results for the impact of mergers on employees.

returns. Using dummy variables for the days of the event period can partly explain the time pattern; however, observations of the same day are likely to be clustered. Hence, one should expect a high degree of within-cluster correlations that bias the OLS estimates of the variance covariance matrix. Applying a modified sandwich estimator avoids the problem of within-cluster correlation and yields robust t-statistics respectively p-values. For that purpose, I construct 31 clusters, one for every day of the event period, and consider every cluster as a kind of super-observation to calculate an unbiased variance covariance matrix. This method is a simple extension of the standard Huber and White robust estimation procedure. Henceforth, the average observation of a cluster is used for deriving the estimates of the variance matrix. Consequently, within-cluster correlations can be neglected – but the clusters have to be independent among each other.

My cross-sectional model for explaining the cumulated abnormal return of each stock in the year 1908 is expanded to work as a panel model in which the abnormal returns of each day are the dependent observations. My model takes, hence, the following form.[109]

$$\begin{aligned}\hat{\varepsilon}^{*}{}_{it} = \beta_0 + \beta_1 \log(cap_i) + \beta_2 \log(age_i) + \beta_3 Success_i + \beta_4 Change_i + \beta_5 Cash_i + \\ \beta_6 DivGrowth_i + \beta_7 Bank_i + \beta_8 Mining_i + \beta_9 Target_i + \beta_{10} Mean_i + \beta_{11} Disclose_i + u_i\end{aligned} \tag{4.2}$$

Running this regression yields the outcome summarized in table 4.1. Caused by heteroscedasticity uncovered by a general White test (p-value: 0.035) and the clustered observations, OLS yields biased estimates for the p-values; thus, the modified robust Huber and White estimator leads to more reliable results.

Before interpreting these results, an autoregression of the residuals ensures that autocorrelation does not exist, otherwise I cannot treat the clusters as independent observations and the modified sandwich procedure is inappropriate. Indeed, table 4.1 stresses that autocorrelation of order two is persistent. Note that it could be useful to include all explanatory variables into the autoregression if an endogenity problem arises – but the results are very close to each other. Obviously, there are two possible ways out, namely a GLS procedure that accounts for the second order autocorrelation and the robust estimation using

[109] Note that the causality problems mentioned in chapter two is considered by inserting the estimated mean returns into this equation to account for the estimation period. It should also be stressed that as long as contemporaneous exogenity conditions hold – this means that exogenity is fulfilled for every of the two equations presented in chapter two – I can stick to OLS estimation without getting biased results. I, thus, concentrate only on the first equation. For a broader discussion of the explanatory variables see chapter two.

the Newey-West[110] estimator. Caused by the problem that I use more cross-sectional than time series dimensions, a GLS estimation is not valid as argued by Beck and Katz (1995). Hence, table 4.1 also contains the outcomes of the Newey-West approach.

Table 4.1: Results of regression (4.2) and autoregressions based on residuals

I run regression (4.2) with p-values based on a modified Huber/White sandwich estimator and a panel Newey-West procedure that controls for autocorrelated residuals.

	OLS estimation with a modified Huber/White method		OLS estimation with Newey-West procedure		Autoregression based on residuals	
Constant	0.0358	(0.907)	0.0358	(0.914)	0.0076	(0.980)
Log(cap_i)	-0.0312	(0.171)	-0.0312	(0.196)	-0.0027	(0.925)
Log(age_i)	0.0180	(0.774)	0.0180	(0.783)	0.0027	(0.962)
$Success_i$	-0.0419	(0.867)	-0.0419	(0.865)	-0.0122	(0.948)
$Change_i$	-0.0178	(0.872)	-0.0178	(0.834)	-0.0659	(0.602)
$Cash_i$	0.0376	(0.514)	0.0376	(0.533)	0.0062	(0.936)
$DivGrowth_i$	-0.0198	(0.769)	-0.0198	(0.777)	0.0152	(0.887)
$Bank_i$	0.1354	(0.049)	0.1354	(0.115)	0.0037	(0.970)
$Mining_i$	0.0787	(0.451)	0.0787	(0.512)	0.0043	(0.974)
$Mean_i$	-1.8998	(0.010)	-1.8998	(0.011)	-0.0957	(0.728)
$Target_i$	0.0193	(0.766)	0.0193	(0.797)	-0.0151	(0.866)
$Disclose_i$	-0.0729	(0.362)	-0.0729	(0.422)	0.0036	(0.970)
Residual (lag 1)	-		-		0.1422	(0.000)
Residual (lag 2)	-		-		-0.0776	(0.005)
Residual (lag 3)	-		-		-0.0109	(0.696)
Adjusted R^2	0.06		0.06		0.01	
F-test (p-value)	9.27	(0.000)	1.48	(0.132)	2.31	(0.004)
Observations	1426		1426		1288	
White test (p-value)	62.52	(0.035)				

The estimated mean return reduces the observed abnormal return significantly (p-value 0.011). This follows directly from the definition of the abnormal return that is the deviation of

[110] A detailed description can be found in Newey and West (1987). To apply this procedure, one has to modify this typical time series approach, which is possible in STATA by executing the `newey2´ command.

the current return from the normal return derived from the estimation period.[111] Unfortunately, this is the only significant coefficient.

4.4.2 How can the time path of abnormal returns be modeled?

The clustering of observations influences the estimated variance-covariance matrix – but if the model is correctly specified and exogenity conditions are fulfilled, the estimated coefficients using OLS are not biased. However, the way the observations are clustered indicate that there remains still another problem – the stability of the regression coefficients over time. Putting this in other words, I seek structural changes in my single equation model. If structural changes are not taken into account, the model is often regarded as misspecified and the OLS coefficients are seen as biased. Since the influential papers of Chow (1960)[112] and Durbin et al. (1975),[113] the discussion about the ways to detect structural changes and whether to define a precise alternative hypothesis[114] stays controversial. Especially, whether uncovered structural changes should be modeled using, for instance, a set of dummy variables, is highly disputed.[115] Looking at the single equation model (4.2), one can argue that the changes in abnormal returns over the days of the event period are not reflected in the included explanatory variables. This variables may be capable to explain the cumulated effect of a merger announcement – but they contribute less to explaining the daily adaptation of stock prices. Thus, inserting dummy variables to model the special influence of the days close to the event day seems to be appropriate. Unfortunately, lacking data on daily trading volume limits the possible alternatives for the year 1908 and makes the usage of dummies attractive. However, which days of the event period are essential for the adaptation process, and should one assume that there are no differences among cross-sectional units?

There are two different opportunities. First, the results with regard to the significance of portfolio-weighted abnormal returns can be used to justify the insertion of dummies for days showing highly significant abnormal returns. Second, one allows different time paths of the adaptation process among cross-sectional units and use stock specific dummy variables for important days. The latter approach obviously has the disadvantage that the power of the

---

[111] In contrast to my former study, I do not build up a simultaneous equation model that is needed to distinguish precisely between direct and indirect effects – but even without a second equation the estimated mean has a high explanatory power and should not be neglected.

[112] Chow (1960) used F-tests to compare a model with and without structural change; thereby, the location of the brake must be known. There exist also approaches that do not require the a priori knowledge of the break points (see Andrews, 1993).

[113] Durbin et al. (1975) cumulated the recursive residuals to obtain a CUSUM test. Based on the common OLS residuals, Krämer and Ploberger (1992) provided an alternative CUSUM approach.

[114] See Perron (1989).

[115] Often one refers to the Lucas (1976) critic when talking about macroeconomic model that should be stable over time.

event-study is diminished because cross-sectional units are no longer aggregated; hence, exogenous time shocks and remarkable events both cause abnormal returns and show similar patterns. Therefore, the possibility to decide whether the abnormal returns are indeed due to the merger announcement or to a time shock is lost. But assuming the same adaptation path is also far from being true, especially recalling the results provided in chapter two and three. I can reduce this problem by aggregating not over the whole sample but by defining subgroups; thereby, dividing the firms that disclose mergers and hidden information seems to be obvious.

Besides the possibility to justify the chosen set of dummy variables by looking at significant daily returns, a more formal method to identify structural changes in a panel data setting should be taken into account. Note that significant abnormal returns are not a perfect indicator for a structural change in my single equation model because parts of significant abnormal returns can be explained. Thus, considering the residuals of equation (4.2) is preferable.

Moreover, using two different explorative methods to detect structural breaks in a panel data setting, namely a recursive CUSUM and an OLS based CUSUM test, provides clear evidence for the usage of dummy variables. Note that applying a simple Chow test[116] to uncover a structural break point is inappropriate when dealing with several structural changes during the event period.

Furthermore, the simple and often sufficient solution to insert 30 time dummies and to test against a reference group does not yield useful results. Regardless which reference day is chosen or whether possible interaction terms are considered, it is never possible to obtain any significant coefficients of dummy variables. Therefore, the next two sections discuss the recursive and OLS based CUSUM test for panel data.

4.4.3 The recursive CUSUM approach

The basic idea behind the recursive CUSUM method is straightforward. Using only the information available at t-1, I estimate the parameter vector $\mathbf{b_{t-1}}$. The recursive residual is then obtained by calculating the difference between the true value of the dependent variable $\mathbf{y_t}$ and its forecasted value based on previous information.[117] Note that the time index t is appropriately chosen for the event period covering 31 days.

$$\mathbf{\tilde{u}_t} = \mathbf{y_t} + \mathbf{X_t b_{t-1}} \qquad \forall t \in \{2,3,...,31\} \tag{4.3}$$

---

[116] Of course, there exist so called F-supremum Chow tests that can be used under this condition (see Andrews, 1993).
[117] This section follows Han and Park (1989).

Note that the cross-sectional dimension n has the value forty-six and the column vectors in (4.3) are, thus, n×1 dimensional. Imposing the assumption that recursive residuals are spherical, the variance-covariance matrix[118] of recursive residuals has the following shape.

$$\mathbf{\Sigma_r} = \frac{1}{31}\sum_{t=1}^{31}\mathbf{\tilde{u}_t'\tilde{u}_t}\left(\mathbf{I_{n\times n}} + \mathbf{X_r}\left(\sum_{\mathbf{s=1}}^{\mathbf{r-1}}\mathbf{X_s'X_s}\right)^{-1}\mathbf{X_r'}\right) \tag{4.4}$$

The first term is just the method of moments estimator for the error term variance that is supposed to be constant over time. Whereas, the second term stems from the restrictive usage of information at time r; hence, there is inaccuracy in the covariance measurement if one relies only on the information available up to time r. Obviously, when r converges to T (=31), the second term converges to the identity matrix with dimension n×n.

Detecting second order autocorrelation as well as within-cluster correlation violates the imposed assumption that recursive residuals are spherical – but the scale of the violation seems not to be so severe to require a more complicated approach.[119]

The resulting recursive residuals from equation (4.3) are standardized using the estimated variance-covariance matrix (4.4) and aggregated over time. In addition, I use two different ways to carry out this test. First, figure 4.3 plots the portfolio weighted recursive residuals against its critical boundaries.[120] Second, by constructing n (=46) individual critical regions, the CUSUM test gives hints regarding structural breaks over time for every cross-sectional unit. To avoid drawing forty-six critical regions, figure 4.4 summarizes the results by plotting the number of cases in which the critical values are exceeded against the day of the event period. Obviously, structural breaks are only observable on the individual level, whereas figure 4.3 shows that no structural breaks occur when one considers the whole panel data set.

---

[118] The derivation of the variance-covariance matrix is based on Han and Park (1989).

[119] Han and Park (1989) also derived a modified approach that accounts for first order autocorrelation in residuals.

[120] Brown et al. (1975) derived the critical values for the univariate case. One can rely on this critical values when using the aggregate CUSUM test.

Figure 4.3: Aggregate recursive CUSUM test for the panel regression (4.2)

The aggregate recursive CUSUM test is carried out; thereby, the resulting empirical fluctuation process and the critical values are plotted below.

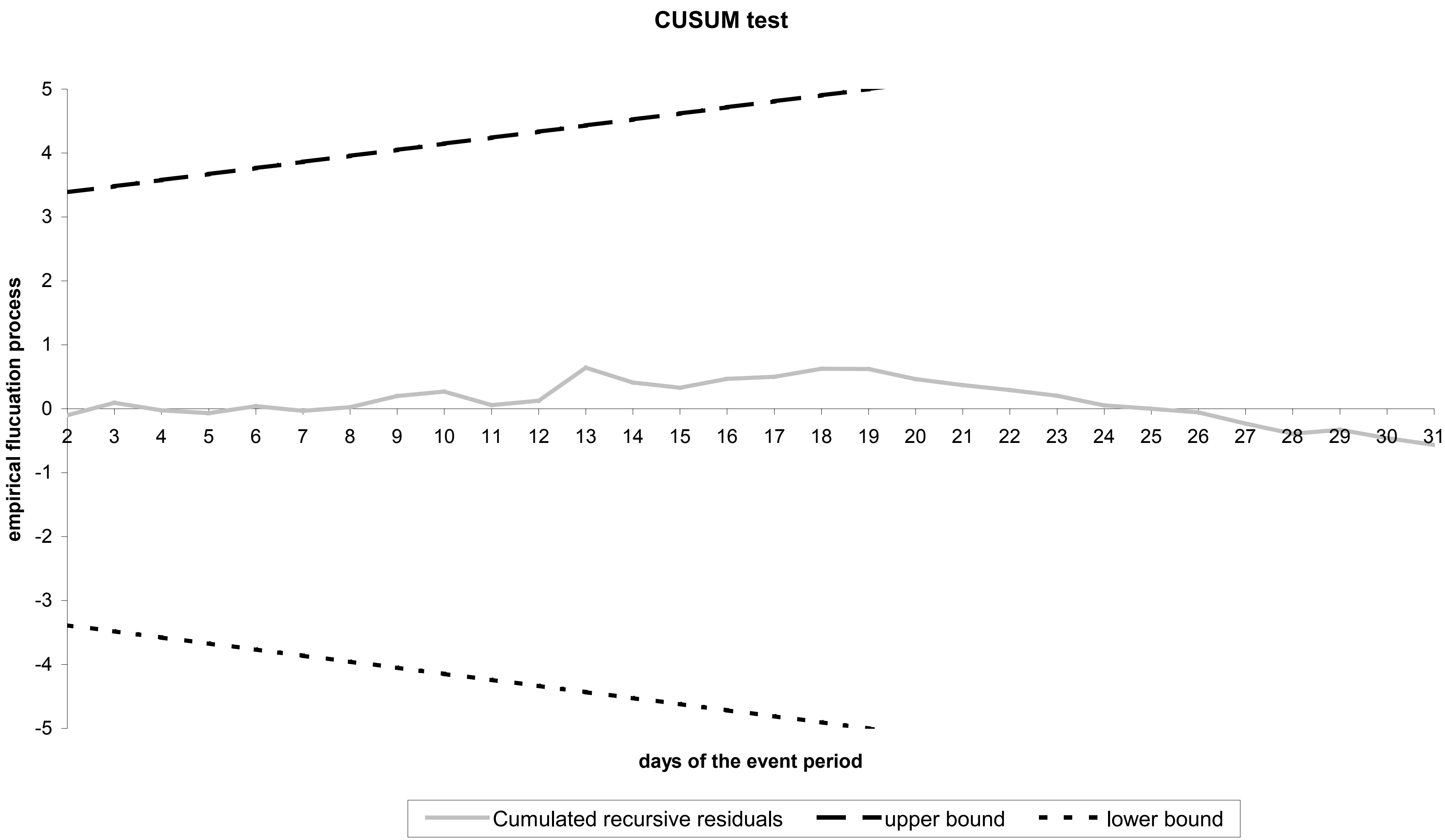
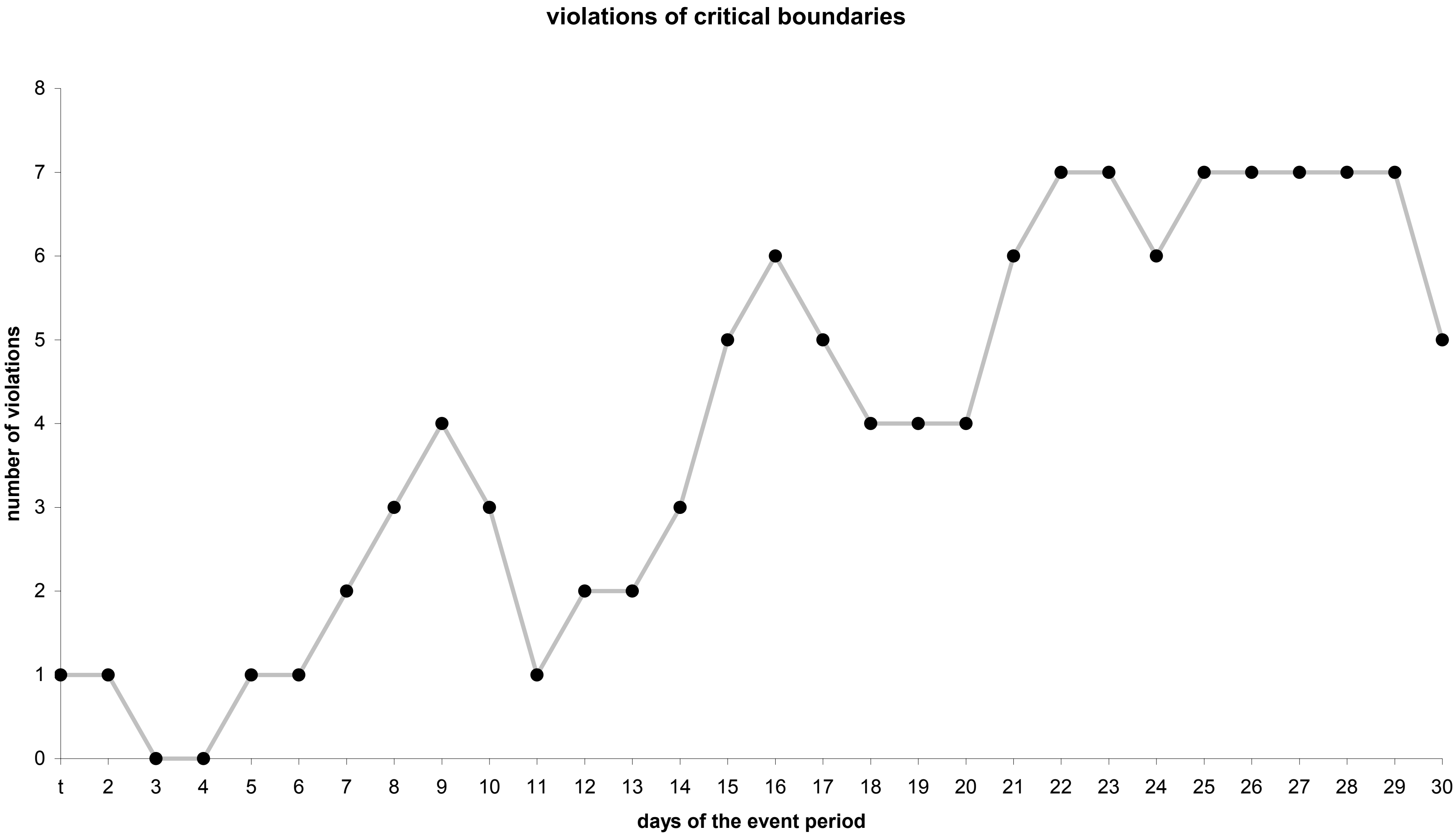


Figure 4.4: Number of structural breaks on the individual level

The graph shows the number of structural breaks – indicated by a recursive CUSUM tests – that occur on each day of the event period.



There is an increasing literature about the problems of the recursive CUSUM tests to detect structural breaks that affect the empirical fluctuation process very early or relatively late duirn the investigation period. Fortunately, there is an alternative that tackles this problem.

4.4.4 The OLS based CUSUM test[121]

Using the methodology developed by Kramer and Ploberger (1992) who propose a linear boundary that is proportional to the standard error of the underlying Brownian motion,[122] figure 4.5 plots the empirical fluctuation process of OLS residuals from regression (4.2). Because this area of research is still a highly active research field, alternative ways to define boundaries exist. Having the advantage to detect early and late structural breaks with higher power, Zeileis (2004) used a circular boundary depicted in figure 4.6. In contrast to the former literature on OLS based CUSUM tests that focused on univariate time series, I aggregate over the cross-sectional units and carry out an aggregate CUSUM test with critical values obtained from the univariate studies. On the aggregate level, structural changes cannot be detected regardless which boundary determines the rejection area.

4.4.5 The failure of the traditional approach

Even sophisticated methods to uncover instable regression coefficients fail to justify the use of a set of dummy variables to characterize the time path of adaptation around the merger announcement. Moreover, separating between different kinds of disclose as done in chapter three and modeling differences regarding the time path of abnormal returns by inserting interaction terms does not yield significant results, after controlling for autocorrelation and heteroscedasticity. The recursive CUSUM as well as the OLS based CUSUM test applied to

[121] I skip the technical details to avoid a too detailed and lengthy discussion.
[122] Obviously, under the null hypothesis, adding up OLS residuals results in a Brownian motion. Hence, a violation of this implication gives a hint with regard to structural breaks in the suggested regression relation.

Figure 4.5: OLS based CUSUM applied to panel regression (4.2) with `traditional´ boundaries

The graph shows the OLS CUSUM denoted as W(t) of regression (4.2); thereby, the boundaries are provided by Kramer and Ploberger (1992).

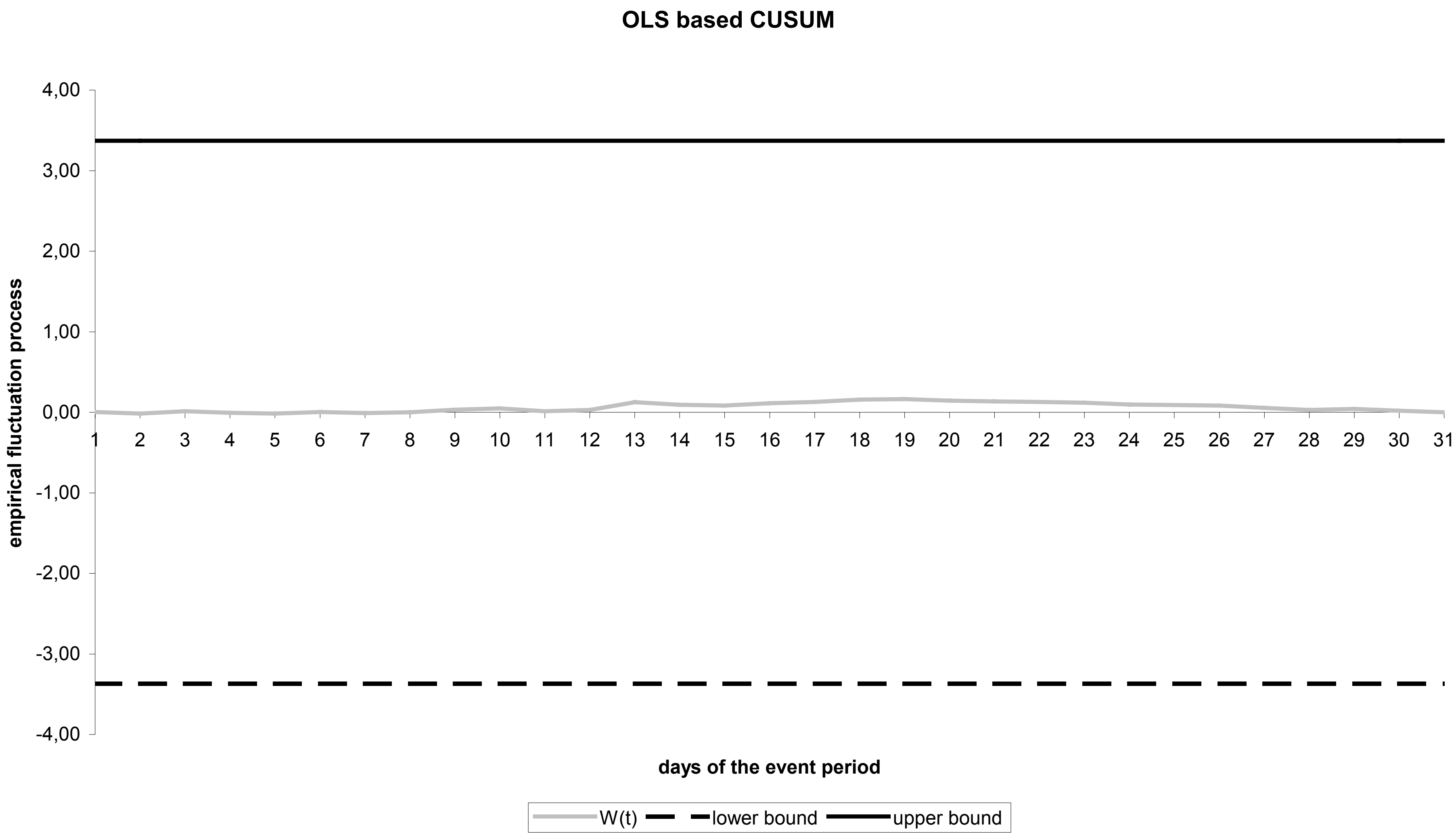
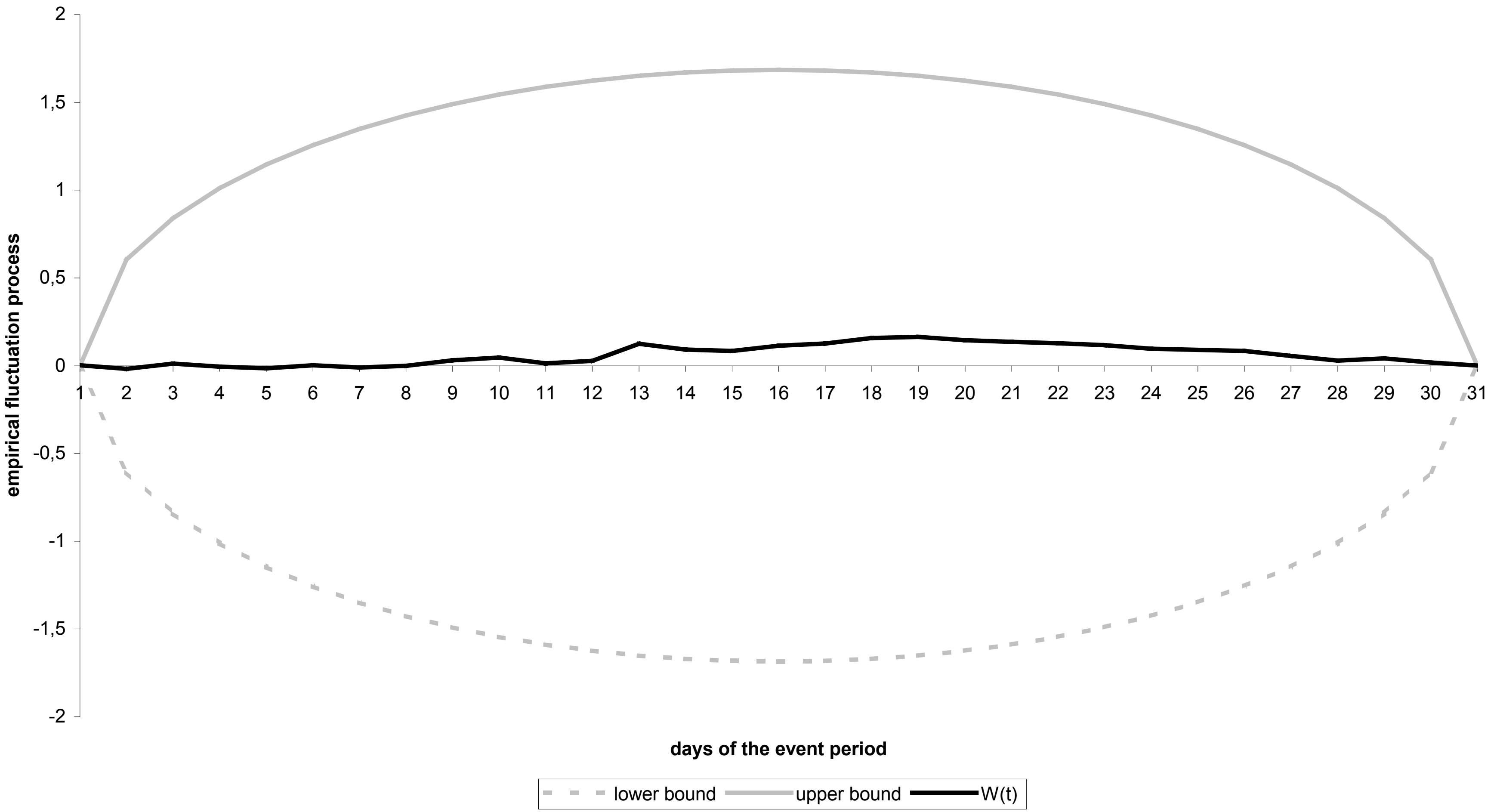


Figure 4.6: OLS based CUSUM with alternative boundaries

The graph depicts the OLS CUSUM denoted as W(t) of regression (1); thereby, the circular boundary proposed by Zeileis (2004) are applied



individual firms indicate structural changes (see figure 4.4) – but this means that one should include company and time specific dummies to capture this changes. Hence, one ends up with a set of dummy variables out of which a dummy stands for one specific abnormal return of firm x at day y. Accordingly, a over-specification seems to be very likely. Of course, the resulting regressions possess a high degree of explanatory power indicated by high adjusted R^2 – but the out-of-sample properties of such regressions are extremely bad. Putting this in other words, one cannot make general statements, for instance, whether firm size affects the adaptation process.

4.5 Explaining the adaptation process of stock prices: A time series approach

4.5.1 Intervention models with transitory shocks[123]

Besides the already mentioned problems of the panel analysis, the time series dynamics of abnormal returns is not exploited to improve the understanding of the adaptation process. Following the Box-Jenkings-approach respectively information criterions, I specify an ARMA(3,3) model for the portfolio weighted abnormal returns. Note that the partial autocorrelation function indicate significance for the third lag. In addition, table 4.2 underlines that an AR(1) specification does not capture the dynamic better than the ARMA(3,3) model. The quality of the respective specification is indicated by the information criterions, namely Akaike and Schwarz, as reported in table 4.2. Therefore, excluding insignificant lags and specifying an AR(1) process is inappropriate. This finding is due to a high degree of collinearity among the ARMA(3,3) explanatory variables that is responsible for the low p-values – but does not say anything about the total explanatory power of the model. For the purpose of forecasting, one is usually more interested in having a model with high explanatory power than evaluating the partial effects of specific lags. Figure 4.7 depicts the portfolio weighted abnormal returns and its fitted values using the ARMA(3,3) model. In addition, the upper and lower bound of a 95% confidence interval indicate that three days before the announcement (t=13) the abnormal returns are unusually high.

[123] A detailed description of intervention models can be found in Mills (1990).

Table 4.2: ARMA specification to capture the dynamics in abnormal returns

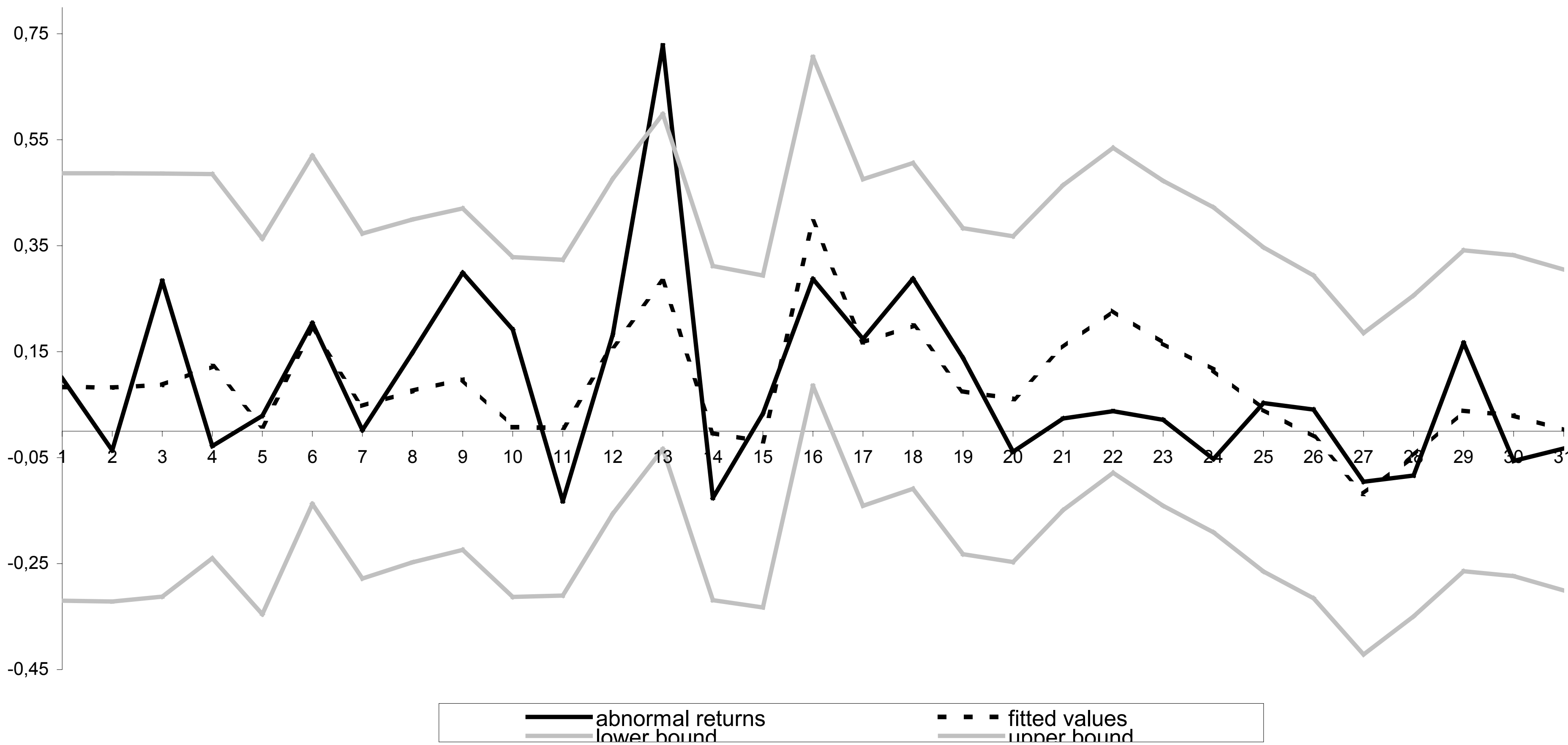
The portfolio weighted abnormal return can be described as ARMA(3,3); thereby, excluding insignificant lags and specifying an AR(1) process does not capture the correct dynamics. Note that p-values are set in parentheses.

	ARMA(3,3)		ARMA(1,0)	
Constant	0.0834	(0.094)	0.0889	(0.006)
AR(1)	0.5694	(0.063)	-0.0508	(0.788)
AR(2)	0.0376	(0.922)	-	
AR(3)	0.0901	(0.762)	-	
MA(1)	0.8704	(0.913)	-	
MA(2)	0.2423	(0.947)	-	
MA(3)	-0.7582	(0.920)	-	
Number of observations	31		31	
Log Likelihood	18.02		10.69	
Akaike criterion	-22.05		-17.38	
Schwarz criterion	-12.01		-14.51	

Of course, an explanatory variable like time can play a role in determine the abnormal returns. Hence, including time as a dummy variable can help to explain the transitory increase of abnormal returns at t=13, three days before the announcement. This is a simple intervention model; thereby, the thirteenth day of the event period takes the value one and all other days have an impact of zero. Inserting the impulse dummy at t=13 into the ARMA(3,3) model yields to a highly significant positive coefficient (0.4657, p-value: 0.000). Comparing the residuals of the ARMA(3,3) model with the residuals of the intervention model shows that the peak at t=13 disappears – but a new peak at t=16, the announcement day, emerge. Hence, one can improve the intervention model by taking into account the transitory shift at t=13 and t=16. It turns out that the transitory shock at t=16 is insignificant (p-value 0.195) and the sequence of the error terms using one respectively two shocks are almost equal. Although

Figure 4.7: ARMA(3,3) representation of portfolio weighted abnormal returns and the 95% confidence interval

This figure plots the process of abnormal returns and fitted values of an ARMA(3,3) representation. Upper and lower boundaries of the 95% confidence interval set the expected range of movement.



such so called intervention models are straightforward to implement, there remain always doubts about the way to specify the correct form of intervention.

4.5.2 Transfer function models

To overcome the inherent reliability problem of intervention models, one can extend the concept of interventions by allowing a non-predefined time paths of shocks. Accordingly, I specify an ARMA(3,3) model with a transfer function γ(L) that depends on the absolute distance from the event day. Hence, it takes the following form.

$$\bar{\varepsilon}_t = c + \sum_{j=1}^{3} a_j \bar{\varepsilon}_{t-j} + \sum_{j=1}^{3} m_j u_{t-j} + \gamma(L)\tau + u_t \qquad (4.5)$$

Where: $\bar{\varepsilon}_t$ ....Portfolio weighted abnormal return at day t of the event period

c ....Constant

a_j ....Coefficients of the autoregression

m_j ....Coefficients of the moving-average component

γ(L) ....arbitrary polynomial in the lag operator L

τ ....Absolute distance of the day t from the event day

u_t ....Contemporaneous error term

The aim of this analysis is to figure out how the distance to the public merger announcement influences the process of abnormal returns; thereby, a specific time path of the impact is not predetermined. Note that the exogenous variable τ is related to the time variable t; this distinguishes my analysis from usual transfer function models. Calculating and interpreting a cross-correlogram between the abnormal returns and the exogenous variable τ uncovers the shape of the transfer function γ(L). Because the exogenous variable τ is defined as distance from the event day, analyzing the cross-correlogram determines the \`empirical announcement day´ which is the day that exhibits the highest impact on the series of abnormal returns. If the public merger announcement is anticipated or insider trading affects prices prior to the revelation, one should expect that the \`empirical announcement day´ occurs before the newspaper makes the information public. Figure 4.8 depicts the empirical cross-correlogram for the whole sample and the upper and lower limits based on two standard deviations. The variable τ possesses a significant cross-correlation with the process of abnormal returns at Lead three. How can one interpret this result? Consider the definition of the variable τ as the absolute difference of day t from the announcement day. If the third lead of τ affects the time series of abnormal returns, the \`empirical announcement day´ is at t=13, three days before the merger appears in the newspaper. This empirical finding underlines the presence of insider activity respectively a high degree of anticipation in the year 1908. One can regard this procedure as an alternative to applying an event-study and measuring the run-ups prior to the official announcement. After determining the appropriate lag respectively lead of the transfer function τ, I can write the ARMA(3,3) with a modified explanatory variable τ_{+3} that measures the absolute difference of day t from the thirteenth day of the event period in the following manner. Table 4.3 summarizes the outcomes.

$$\bar{\varepsilon}_t = c + \sum_{j=1}^{3} a_j \bar{\varepsilon}_{t-j} + \sum_{j=1}^{3} m_j u_{t-j} + \gamma\tau_{+3} + u_t \tag{4.6}$$

Table 4.3: Outcomes of the transfer function model (4.6)

Note that multi-collinearity is high among AR and MA terms. P-values are set in parentheses.

	ARMA(3,3) with transfer function	
Constant	0.2031	(0.000)
AR(1)	-0.1676	(0.868)
AR(2)	-0.5635	(0.579)
AR(3)	-0.2034	(0.841)
MA(1)	0.2797	(0.774)
MA(2)	0.2618	(0.512)
MA(3)	-0.3797	(0.308)
τ_{+3}	-0.0143	(0.000)
Number of observations	31	
Log Likelihood	21.07	
Akaike criterion	-26.15	
Schwarz criterion	-14.67	

To illustrate the intuition behind equation (4.6), figure 4.9 separates the deterministic component as determined by the transfer function and the AR(3) process from the stochastic component, the ARMA(3,3) model. Note that the transfer function is influenced by the autoregressive nature of the time series.

This illustrates one of the undisputable advantages of transfer function models in comparison to my former panel analysis; it is possible to divide the movement into a stochastic component and a deterministic component. The following two sections discuss the application of the transfer function analysis to the disclosure problem of the year 1908 and the speculation of the year 2000. The last section, then, emphasizes the pitfalls of this method.

Figure 4.8: Cross-correlation between portfolio weighted abnormal returns and the absolute distance from the announcement day

This figure depicts the cross-correlation between the abnormal returns and different lags or leads of the variable τ

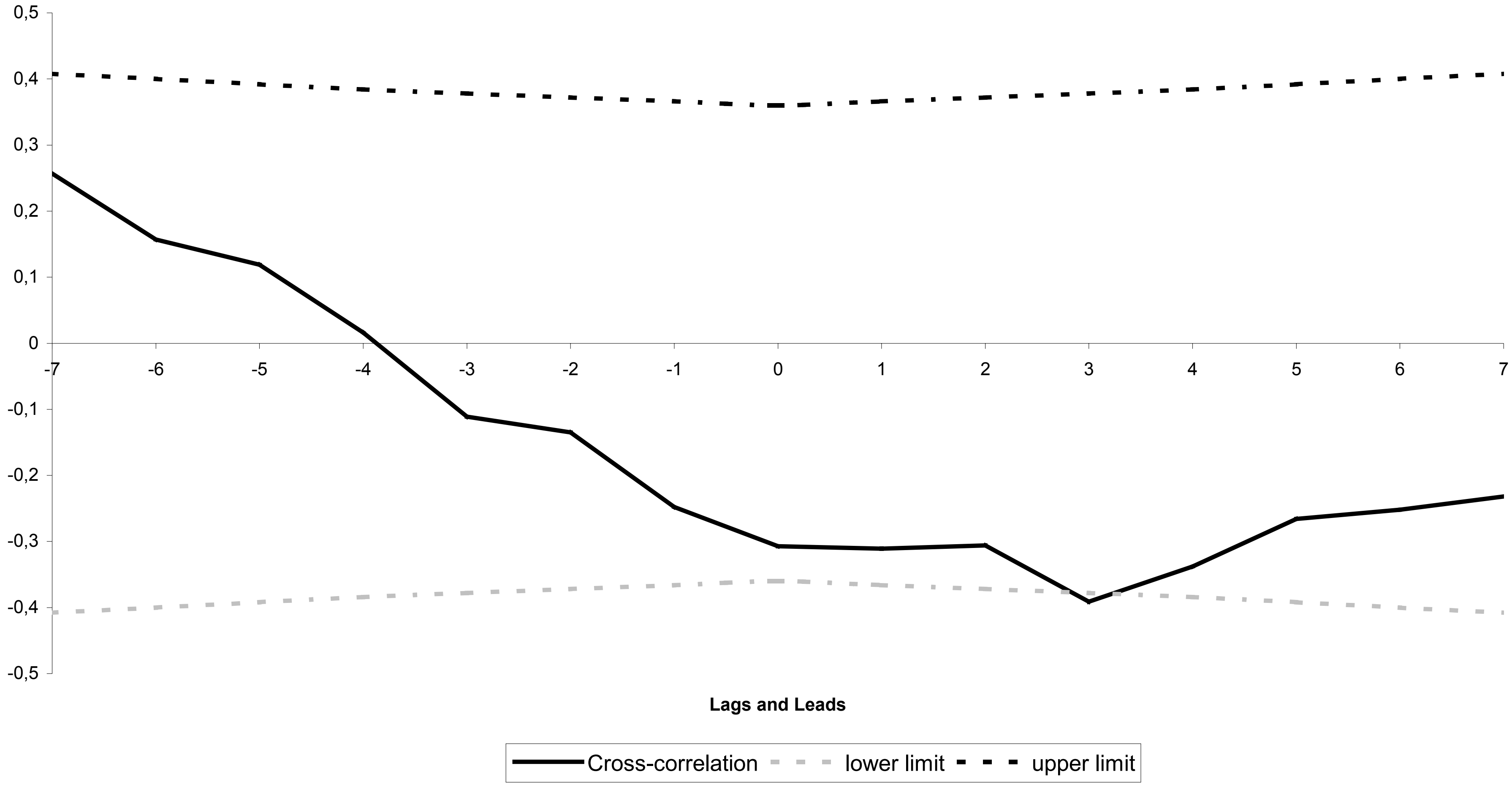
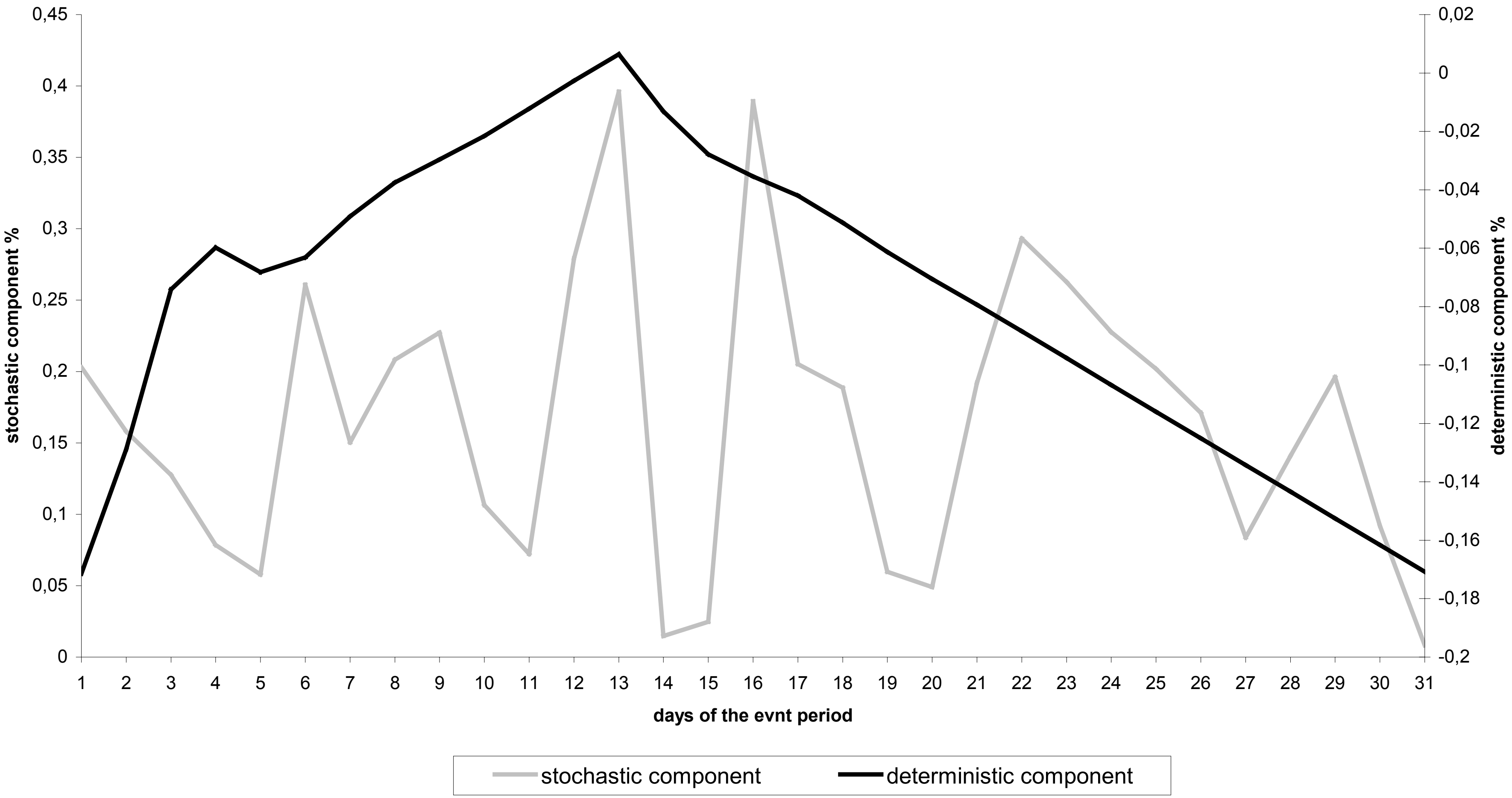


Figure 4.9: Deterministic component due to the announcement and stochastic fluctuations in abnormal returns

Figure 4.9 separates the deterministic component as determined by the transfer function from the stochastic component, the ARMA(3,3) model



4.5.3 Different ways of disclosure and the `empirical announcement day´

The idea of the following section is to determine the `empirical event day´ in the case of hidden and disclosed mergers. Using the same approach as in the previous section, I conclude that no significant cross-correlation between the abnormal returns and the parameter τ can be found if the firms disclose mergers. This means that the announcement day possesses no extraordinary effect on the process of abnormal returns – but all other days of the event period share the same feature. This finding stems from the fact that a large portion of the fluctuation in abnormal returns is already captured in the ARMA(3,3) specification. In contrast, I detect significant cross-correlation at the leads one, two, and three if firms tried to hide mergers. Hence, the following transfer function model describes the behavior of abnormal returns, and table 4.4 contains the results.

$$\bar{\varepsilon}_t = c + \sum_{j=1}^{3} a_j \bar{\varepsilon}_{t-j} + \sum_{j=1}^{3} m_j u_{t-j} + \gamma_{+1}\tau_{+1} + \gamma_{+2}\tau_{+2} + \gamma_{+3}\tau_{+3} + u_t \tag{4.7}$$

Of course, the three empirical event days, one to three days before the public announcement, exhibit overlapping deterministic influences on the process of abnormal returns. Moreover, two coefficients are negative, whereas one is positive. In general, one can conclude that the empirical event day is prior to the official merger announcement respectively the published rumor if the firm tries to hide information.

Now, to improve the model fit further, one can change the MA(3) specification because using the transfer function reduces the need for a moving-average representation of the residuals. Inspiring the ACF (autocorrelation correlation function) plot after inserting the transfer function makes a MA(3) component redundant because the series does not indicate significant autocorrelations at any lag. Hence, one can estimate a more comprehensive model that has a better fit to the data. Table 4.4 contains this alternative specification as well.

$$\bar{\varepsilon}_t = c + \sum_{j=1}^{3} a_j \bar{\varepsilon}_{t-j} + \gamma_{+1}\tau_{+1} + \gamma_{+2}\tau_{+2} + \gamma_{+3}\tau_{+3} + u_t \tag{4.8}$$

Table 4.4: Transfer function analysis for undisclosed mergers in the year 1908

Specifying the abnormal returns as ARMA(3,3) respectively ARMA(3,0) and using the three days before the public merger announcement as empirical event days gives the following outcomes. Note that multi-collinearity is very high among AR and MA terms.

	ARMA(3,3) with transfer function		ARMA(3,0) with transfer function	
Constant	0.6976	(0.000)	0.6913	(0.001)
AR(1)	0.1034	(0.863)	0.5488	(0.029)
AR(2)	0.0901	(0.877)	-0.4648	(0.052)
AR(3)	-0.1382	(0.756)	0.0529	(0.804)
MA(1)	-0.2223	(0.958)	-	
MA(2)	0.6853	(0.893)	-	
MA(3)	0.5201	(0.820)	-	
τ_{+1}	-0.6929	(0.001)	-0.6725	(0.001)
τ_{+2}	1.3445	(0.001)	1.3490	(0.001)
τ_{+3}	-0.7131	(0.000)	-0.7349	(0.000)
Number of observations	31		31	
Log Likelihood	-9.93		-13.48	
Akaike criterion	39.85		40.97	
Schwarz criterion	54.19		51.01	

4.5.4 Does speculation affect the `empirical announcement day´?

The sequence of portfolio weighted abnormal return of the whole sample, consisting of 61 cases observed in 2000, can be best described by an ARMA(1,1) specification because the ACF and PACF plot indicate significance for the first lags and information criterions confirm the specification. If I stick to the definition of the variable τ, the cross-correlogram fails to uncover significant cross-correlations regardless which lead or lag is taken into consideration. Therefore, the use of a linear transfer function does not fit to the data of the year 2000.

4.5.5 The limitations of the transfer function analysis

Despite their fascination as an analytical tool to figure out the `empirical announcement day´, my modified version of a transfer function model is accompanied by several problems. Maybe the most disputable part is the determination of the correct ARMA specification of the underlying process of abnormal returns. Interpreting ACF and PACF plots has more to do with art than with pure science – some critics argue. Nevertheless, using information criterions is also sometimes ambiguous. For instance, following the Akaike criterion, one should prefer an AR(3) specification; however, the Schwarz criterion favors an ARMA(3,3) model as shown in table 4.4.

Besides this general specification problem, an additional issue arises: How should one predetermine the exogenous variable? Note that I obviously cannot collect data of the exogenous variable τ; I have to define τ. But with the definition of τ as absolute difference between the current day t and the event day (t=16), the linear shape and a single peak of the estimated transfer function is also given. Hence, changing the definition of τ, for instance measuring the quadratic deviation from the event day, also affects the transfer function and the deterministic component. So to compare different subgroups, for instance hidden and disclosed mergers, I recommend to stick to the determined ARMA model as well as to the definition of the exogenous variable τ. Consequently, my result that given the definition of τ hidden mergers exhibit an empirical event day prior to the newspaper announcement in comparison to disclosed mergers that possess no empirical event day stays valid. Nevertheless, one can argue that caused by a higher degree of informational efficiency in the year 2000, the transfer function should not have a linear shape. Instead, the exogenous variable can be defined as the square root of deviation from the event day. Using this definition, the cross-correlogram of the year 2000 shows significant coefficients for the leads two and three. Thus, the irrational speculation that starts about four to three days before the merger announcement affects the `empirical event day´. The transfer function model states that the empirical event occurs three days earlier than the newspaper announcement. Of course, the modified transfer model cannot distinguish between irrational speculation and insider activities, as I did in my former study presented in chapter three. Note that the often discussed problem that the explanatory time series must be exogenous – to be more precise, the model require strict exogenity – does not occur in my setting because the explanatory variable follows from a definition. In addition, the error term after imbedding the transfer function is close to a white-noise process; hence, the sequence of residuals is uncorrelated with τ regardless which lag is tested.

Weighing up the pros and cons, I conclude that modified transfer function analysis is an interesting extension of my former studies – but an event-study approach provides even more insights into the short term effects of merger announcements.

4.6 Event-induced uncertainty in daily abnormal returns

4.6.1 Developing a basic panel based GARCH approach for abnormal returns

Engle's (1982) seminal paper paved the ground for an increasing analytical interest regarding the conditional variance structure of economic time series. His concept that allows an autoregressive structure in squared residuals (ARCH) was later extended by Bollerslev (1986)[124] to capture in addition a moving-average structure (GARCH). The standard procedure starts with an ARIMA specification of the initial time series. The ACF and PACF plots of the resulting squared residuals indicate the appropriate GARCH specification. Besides ARIMA models, a multiple regression in which the initial time series serve as dependent variable are the starting point of an GARCH analysis.

In contrast to these standard procedures, my GARCH model is applied to the time series of abnormal returns. This series shares common features with the above mentioned residuals of regressions or ARIMA models as abnormal returns are deviations of observed daily returns from their normal levels. To detect event-induced uncertainty measured by an increase of volatility around the merger announcement, I estimate an ARCH model for the abnormal returns during the event period. Because the time series of portfolio weighted abnormal returns is relatively short embracing only 31 observations, I estimate a panel based GARCH model.[125] Accordingly, my basic panel data model allows for conditional heteroscedasticity in abnormal returns.

$$R_{it} = \hat{\mu}_i + \varepsilon_{it}^* \quad \forall t \in \{1,2,...,T\}; i \in \{1,2,...,n\} \tag{4.9}$$

$$\varepsilon_{it}^* = u_{it}\sqrt{h_{it}}, \quad \boldsymbol{\varepsilon}_{\mathbf{t}}^* \sim \mathbf{N}\left(\mathbf{0}, \mathbf{I}_{n\times n}\boldsymbol{\sigma}_e^2 + \mathbf{I}_{n\times n}\mathbf{Var}(\hat{\boldsymbol{\mu}})\right) \tag{4.10}$$

$$h_{it} = \alpha + A(L,\boldsymbol{\beta})\varepsilon_{it}^{*2} + B(L,\boldsymbol{\gamma})h_{it} \tag{4.11}$$

If I ascribe the statistical properties derived in chapter two to abnormal returns, I have to write the GARCH(p, q) process in this manner. Obviously, this representation could be directly estimated applying the standard GARCH estimator. However, standardizing residuals would facilitate the model because the unconditional variance should not differ in the cross-section after standardization. Note that the depicted variance covariance matrix of ε_{it} is the unconditional distribution of abnormal returns as derived from the CMR model. Thereby, σ_e^2 is the n×1 dimensional vector of the error term's variance resulting from the CMR model

[124] Bollerslev (1986) showed that by using a GARCH(1,1) specification the need for ARCH models with several lags is no longer given.

[125] Cermeño and Grier (2001) discussed the superiority of the GARCH estimator in comparison to the OLS estimator in a panel setting.

during the estimation period[126] and the Var(μ) represents the n×1 dimensional vector of the variance of the estimated sample average μ. To estimate the GARCH(p, q) model, I favor to standardize the abnormal returns in the following way.

$$\begin{aligned} \mathbf{\Omega}_{\varepsilon} &= \mathbf{I}_{n \times n} \boldsymbol{\sigma}_{e}^{2} + \mathbf{I}_{n \times n} \mathbf{Var}(\hat{\boldsymbol{\mu}}) \\ \mathbf{pp}' &= \mathbf{\Omega}_{\varepsilon} \\ \mathbf{r_t} &= \mathbf{p}' \boldsymbol{\varepsilon}_{\mathbf{t}}^{*} \end{aligned} \tag{4.12}$$

Using standardized abnormal returns r_{it}, the basic GARCH(p, q) model can be rewritten; thereby, the error term z_{it} is standard normally distributed.

$$r_{it} = z_{it} \sqrt{h_{it}}, \quad z_{\mathbf{t}} \sim \mathbf{N}(\mathbf{0}, \mathbf{I}_{n \times n}) \; \forall t \in \{1,2,...,T\} \tag{4.13}$$

$$h_{it} = \alpha + A(L, \boldsymbol{\beta}) r_{it}^{2} + B(L, \boldsymbol{\gamma}) h_{it} \tag{4.14}$$

Note that A(.) and B(.) are polynomial lag operators with the coefficient vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, which obviously have the dimension p×1 and q×1.[127]

4.6.2 Specifying the correct GARCH(p, q) model

Besides inspiring the ACF and the PACF plots,[128] a modified LM approach[129] is common to detect the specification of an ARCH(p) model. A practical guide for an optimal choice of the GARCH dimensions is to detect the maximum lag p that is accepted by the LM test. Then one should use ARCH(1) or ARCH(2) if the maximum number of lags is one respectively two. For higher orders, it is generally better to estimate a GARCH(1,1) model[130] that exhibits a very similar pattern when compared to ARCH processes of higher order. The LM test for panel data takes the following form.

$$r_{it}^{2} = \alpha + \sum_{j=1}^{p} \beta_{it-j} r_{it-j}^{2} + \eta_{it} \tag{4.15}$$

Thus, the squared abnormal returns of firm i are regressed on their lagged values up to lag p. If the regression possesses a high explanatory power as measured by R^2, the null hypothesis that there is no autoregressive conditional heteroscedasticity of lag p can be rejected with confidence. Thereby, the test statistic nTR^2 is asymptotically Chi^2-distributed with p degrees

---

[126] The error term is just the deviation of the current daily return from the sample average during the estimation period.

[127] The stationarity of this panel based GARCH model is achieved if A(1)+B(1) is smaller than one. This requirement also guaranties that the GARCH process is stationary for every cross-section (see Bollerslev, 1986).

[128] Note that I work with panel data; hence, one should take this into account before deriving the ACF and PACF plot.

[129] Engle (1982) provided a LM test for time series data – but after a slight modification, it can also be applied to panel data.

[130] Bollerslev (1986) recommended that a GARCH(1,1) process is more appropriate than a ARCH(p) process if p becomes larger than two.

of freedom. Note that n denotes the number of cross-sectional units, whereas T stands for the time dimension. Tables 4.5 reports the test statistic for several reasonable lag specifications. This test statistic underlines that there is no autoregressive conditional heteroscedasticity in the series of abnormal returns – but regression (4.15) neglects parameter heterogeneity. To account for this panel specific problem, I allow for different constant terms α among cross-sectional units. These modification enables to run an unbiased LM test; table 4.5 also contains these outcomes.

Table 4.5: Panel based LM test to determine the GARCH(p, q) specification

I carried out a panel based LM test to identify the maximum lag of an ARCH specification

LM test	with constant intercept			
Maximum lag p	R^2	Observations nT	Test statistic	p-value
1	0.0003	1380	0.41	0.522
2	0.0003	1334	0.40	0.819
3	0.0012	1288	1.55	0.671
4	0.0017	1242	2.11	0.716
5	0.0049	1196	5.86	0.320
6	0.0041	1150	4.72	0.580
7	0.0038	1104	4.20	0.756
LM test	with firm specific intercept			
Maximum lag p	R^2	Observations nT	Test statistic	p-value
1	0.0458	1380	63.20	0.000
2	0.0492	1334	65.63	0.000
3	0.0473	1288	60.92	0.000
4	0.0494	1242	61.35	0.000
5	0.0516	1196	61.71	0.000
6	0.0564	1150	64.86	0.000
7	0.0623	1104	68.78	0.000

One can argue that after standardizing the abnormal returns different intercepts α should not be necessary because I control for differing variances across firms. Nevertheless, over long periods the unconditional variance of firm i's return could vary; hence, taking into account the estimated variance during the estimation period by standardizing abnormal returns is not enough. The LM test with firm specific intercepts indicates that even lags larger than seven contribute to explain the squared standardized abnormal returns. Following the standard methodology of this strand of literature, one should specify a GARCH(1,1) model.

4.6.3 The GARCH(1,1) model with and without stock specific effects

Caused by the results of the LM tests that stock specific unconditional variances play a role, one can also incorporate these differences inserting intercept dummy variables into equation (4.14). Unfortunately, using too many dummy variables would cause problems regarding the maximum likelihood maximization procedure. Accordingly, I only include dummy variables that turn out to be relevant when regressing (4.15). On the 1% level of significance and regardless which lag structure is chosen, the `Magdeburger Privatbank´ exhibits a strong increase in unconditional variance from the estimation to the event period. The `Osnabrücker Bank´ is also an exception – but the p-value of the coefficient reaches only 0.076 in regression (4.15). All other companies show no significant differences in unconditional variances over time. Table 4.6 contains two GARCH outputs; thereby, one model accounts for different variances. How can one interpret these results?

In two cases, the series of abnormal returns exhibit a higher than expected unconditional variance around merger announcements. Note that I standardized the abnormal returns such that these residuals should follow a white-noise process with variance equal to one. Therefore, I detect an upsurge in unconditional variances – albeit restricted to only two out of forty-six cases.[131] Consequently, the violations of critical assumptions to derive the test statistics as mentioned in chapter two seem to be less severe – but a recent study due to Savickas (2003) mentioned another problem inherent with event-induced volatility. After estimating my GARCH model, I put some emphasis on this issue.

[131] Obviously, such an increase in unconditional variance biases my test statistics derived in my former study – but controlling for variance increases is hardly fruitful in an event-study setting. Collins and Dent (1984) proposed an GLS approach that tackle this problem; however, in current research in the field of event-studies the GLS model is never used. One problem is the need for lots of observations to estimate a more complicated variance-covariance matrix than just a multiple of the identity matrix.

Table 4.6: GARCH model with and without company specific unconditional variances

Estimating a GARCH(1,1) model with and without stock specific variance yields different estimated parameters.

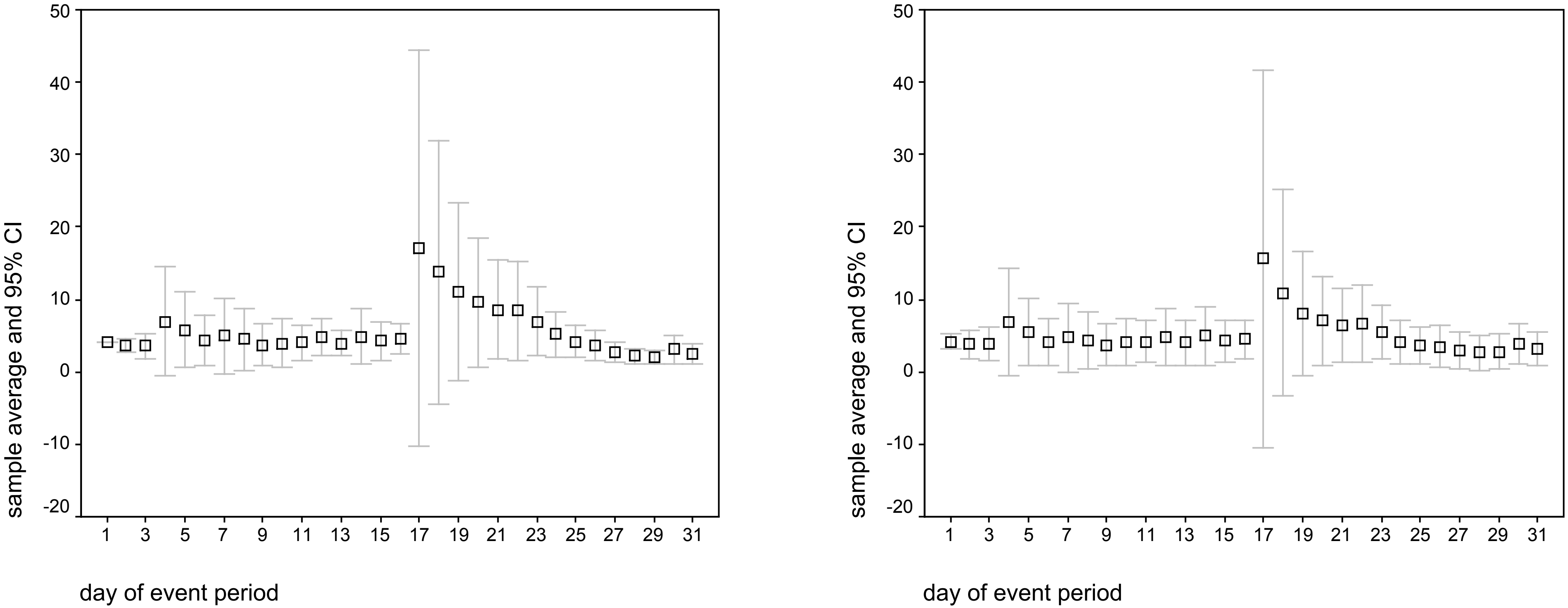
	GARCH(1,1)		GARCH(1,1) with individual variances	
Explanatory variable	Coefficient	p-value	Coefficient	p-value
ARCH1	0.5058	0.000	0.4478	0.000
GARCH1	0.6618	0.000	0.5056	0.000
\`Magdeburger Privatbank´	-	-	3.6261	0.000
\`Osnabrücker Bank´	-	-	4.5006	0.000
Constant	0.1262	0.000	-1.3353	0.000
Log likelihood	-2433.68		-2.314.28	
Observations	1426		1426	

To illustrate the estimated GARCH(1,1) models, I carry out a prediction regarding the conditional variance of returns during the 31 days of the event period. Because I deal with panel data, calculating the sample average of the predicted conditional variances accompanied by 95% confidence intervals may illustrate the results best. Figure 4.10 depicts the predicted conditional variance; thereby the differences between the GARCH(1,1) with or without individual unconditional variances is negligible. Noteworthy, one day after the announcement (t=17), the predicted conditional variance reaches its peak. Thereafter, one can observe a rapid decline. Accordingly, this empirical finding underlines the importance of the newspaper announcements on the degree of conditional volatility.

The last step of an ARCH/GARCH analysis is to make sure that the squared normalized residuals of the model do not exhibit any remaining autocorrelation pattern. Using the procedure (4.15), even an ARCH(1) process can be rejected; the p-value reaches 0.495. Henceforth, my GARCH(1,1) model captures all relevant information regarding the conditional variance.

Figure 4.10: Predicted conditional variance during the event period in the year 1908

In panel A, the GARCH(1,1) model without controlling for individual difference in unconditional variances is used, whereas in panel B individual effects are considered. In both cases, I calculate the sample average of the predicted conditional variances at a specific day t of the event window. To obtain an impression regarding the distribution of these predictions, I also plot the 95% confidence interval for the sample averages.



Panel A **Panel B**

How should one interpret this finding, and can event-induced volatility bias my test statistics? Obviously, newspaper announcements affect not only stock returns but also the volatility of returns. This underlines that new information spread by newspapers has to be reflected in current market values. Based on their information up to the event day (t=16), market participants expect a high degree of volatility on the following day; however, the predicted variance declines rapidly after the announcement. Henceforth, my GARCH approach confirms that merger announcements convey valuable information for the market. Due to the fast decline in predicted volatility after the event day, the high speed with which new information is incorporated into market prices can be also supported.

Despite confirming my former findings based on event-studies, my derived test statistics could be affects by event-induced volatility. Savickas (2003)[132] proposed to divide the standardized abnormal returns r_{it} or standardized cumulated abnormal returns by the square root of expression (4.14). This adjustment corrects for the observed upsurge in volatility around the event day. For my sample, I confirm that the strong increase in volatility affects the p-values reported in chapter two and three for the period from one day to three days after the announcement. Nevertheless, as shown in figure 4.10, the volatility exhibits a pronounced decline about three days after the event day; hence, the distortion of test statistics mitigates rapidly. Discussing the merger paradox, one has to evaluate the total change in market value; hence, the conditional variance at the end of the event window is relevant. Accordingly, the bias seems negligible for analyzing the merger paradox.

4.7 Non-synchronous trading and information from stock price jumps

4.7.1 How important is non-synchronous trading?

The closing prices I used in my studies are not evenly spaced because assets are traded with different frequencies over a trading day. It is also possible that an asset is not traded at all on certain days. Fortunately, the daily newspaper reports \`Brief´ (bid prices) and \`Geld´ (ask prices) that are added to the closing price if at this price only stocks are offered but nobody wants to buy or vise versa. This gives information about the trading patterns without requiring access to daily trading volumes, which are not reported. In general, non-synchronous trading causes many biases like spurious autocorrelation among daily returns.

---

132 Savickas (2003) claimed, based on simulation studies, that his approach is superior compared to the well known non-parametric test statistics provided by Corrado (1989).

For an appraising look, I refer to the results of Lo and MacKinlay (1990)[133] who supposed that periods that show no-trading behavior occur randomly. Thus, strategic interactions that stem from insider trading are not considered. Using my information about bid and ask prices, days without transactions can be identified directly. The model for transaction returns r_{it}^{trans} that are due to executed orders can be written in the following manner.

$$\mathbf{r_t} = \boldsymbol{\mu} + \mathbf{e_t} \qquad \forall t \in \{1,2,...,L\} \tag{4.16}$$

$$r_{it}^{trans} = \begin{cases} 0 & \text{if no trade in t with prob.} \quad \pi_{it} \\ \sum_{j=k}^{t} r_{ij}, k < t & \text{if trade in t with prob.} \quad (\pi_{\text{it}})^{t-k}(1-\pi_{\text{it}}) \end{cases} \tag{4.17}$$

Obviously, equation (4.16) describes my basic CMR model in a log-linear version; thereby, one suggests that the resulting daily returns r_{ij} are not always determined by a transaction. The transaction return r_{it}^{trans} may be zero in period t if no trade occurs, and the newspaper, thus, adds a bid or ask price symbol to the closing price. For illustration, say that over two periods (t=0 and t=1) no trade occurred, and, then, a transaction follows. The resulting transaction return for the last transaction at t=2 is the product of the returns at t=0, t=1, and t=2. Note that the model allows differences among cross-sectional units regarding liquidity.[134]

What happens if a stock is seldom traded? The sequence of transaction returns exhibit higher fluctuations and jumps than does the series of reported returns. Putting this in other words, it states that a rarely traded security may react with time lags due to newly available information – but when it reacts, the reaction is very pronounced.

What effect has non-synchronous trading on the properties of observed returns for individual securities? The most important fact is that the mean of the returns is unbiased by permitting days without trades.[135] This are good news for my event-study approach because the estimated normal return applying the CMR model remains unchanged in the presence of non-synchronous trading. But periods without trades influence the variance as well as the covariance structure of transaction returns; hence, distortions of my former models are possible. Correspondingly, serial correlation could diminish the power of the derived test statistics for the event-study. It is worthwhile, thus, to evaluate the scale of distortion. Lo and MacKinlay (1990) showed that the extent of biases with regard to the variance and the

---

[133] I modify this model with respect to the data generating process of non-observed returns. Because the proposed one-factor log-linear model as used in Lo and MacKinlay (1990) requires the definition of a market index, I stick to my CMR specification. But in contrast to my previous analyses, it is now favorable to use log returns. This simplifies the model.

[134] π_{it} can vary among stocks i.

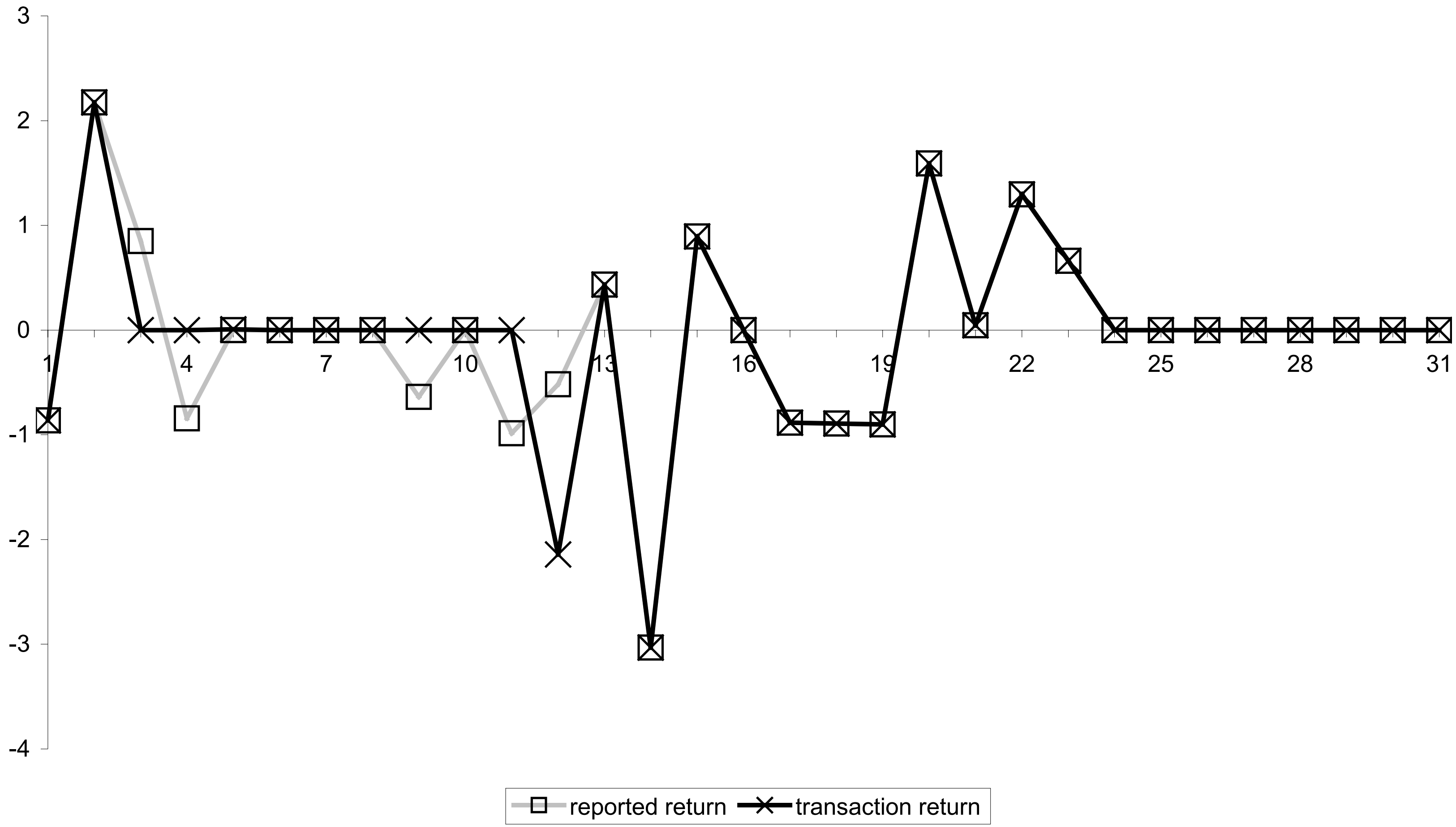
[135] This result can be easily verified by taking the unconditional expectation of equation (4.17). Details can be found in Lo and MacKinlay (1990).

autocorrelation structure is related to the squared mean return. If the mean return is low, the mean-reversion of transaction returns does not cause remarkable distortions of the variance-covariance matrix of returns. Because the mean returns are all close to zero (see chapter two), the maximal bias seems to be negligible.

To illustrate how the transaction return as defined by expression (4.17) and returns based on reported daily closing prices deviate, figure 4.11 presents the two series of returns during the event period for the company \`Bernburger Maschinenfabrik´. This stocks were rarely traded.

Figure 4.11: Non-synchronous trading pattern of the company `Bernburger Maschinenfabrik´

The return as reported in the newspaper is plotted against the return that results from an executed transaction. Note that this stock is an extreme case for illiquidity in my sample.



The stocks of `Bernburger Maschinenfabrik´ were not traded in ten out of thirty-one working days; thus, one can regard this asset as less liquid. The sequences show the expected patterns: once, a transaction occurs the transaction return exhibits high jumps – but is equal to zero if no trade is conducted. Because the influence of non-synchronous trading is even in historical time periods weak, my former results remain unaffected.

4.7.2 What can we learn from the trading patterns?

In spite of the minor effect of non-synchronous trading as discussed in the previous section, one can learn from periods of trading and non-trading about the liquidity and information asymmetry on the market. Because I observed whether a stock price stems from an executed order or if its based only on demand respectively supply, I can `construct´ the bid and ask side of the market. Accordingly, the published bid and offer prices serve as an indicator for the willingness to pay or sell stocks. To get a continuous line, the former bid and offer prices are used untill new prices appear in the newspaper. This gives an impression about the situation on the market around remarkable events.

In the spirit of a case study, inspiring figure 4.12 depicts the constructed ask and bid prices of the `Schleswiger Bank´ during the event period. Note that this firm also belongs to the group of undisclosed mergers; hence, table 4.7 summarizes the newspaper articles as well as an extract from the `Handbuch der deutschen Aktiengesellschaften´ that offers annual reports. Demand steadily exceeds supply before the announcement on 2nd January. Then, stock prices jumped accompanied by a strong fluctuation in prices. In this illiquid market, the first transaction is possible four days before the announcement. This unusual time pattern also stresses that this information broadcasted on 2nd January was new for the public and triggered relative strong stock price movements. Note that the merger was already accepted in September 1907 - but it was not possible to detect an official announcement during this time. One explanation for this finding is that only a few shareholders were involved; hence, the public needs not be informed.

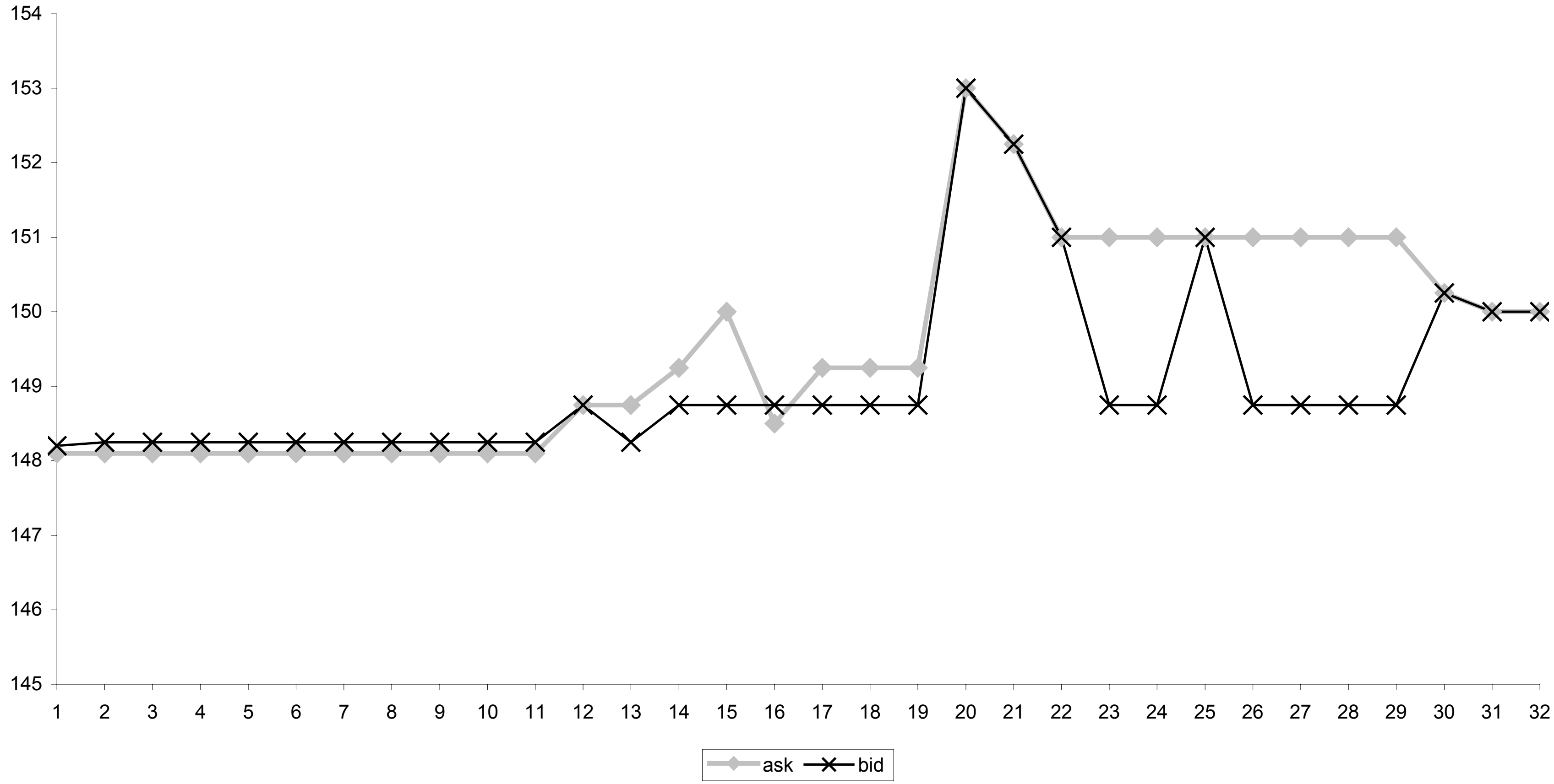
Table 4.7: The case of `Holsten Bank´ and `Schleswiger Bank´

Acquiring firm: `Holsten Bank´ Target firm: `Schleswiger Bank´	
"Berliner Börsenzeitung" – daily newspaper	"Handbuch der deutschen AGs" – annual data
2nd January 1908: `Holstenbank´ announced takeover of `Schleswiger Bank´ **17th February 1908: Page 11; evening issue** annual report of the `Holstenbank´ is published; takeover of `Schleswiger Bank´ executed on 1st January 1908; takeover was already accepted on 9th September 1907 **19th February 1908: Page 14; evening issue** it is made public that already on 13th August 1907 the shareholder gathering of the `Schleswiger Bank´ accepted the merger	**Extract from the annual report:** 9th September 1907: 2000 new shares were issued; thereby, some shareholders of the `Schleswiger Bank´ received 298 shares. **Additional information:** Total volume of this takeover 327800 Mark

Accordingly, the acquirer `Holsten Bank´ could communicate easily with a few major shareholders without using a newspaper statement. In addition, a merger has to be announced in the annual report of the company together with a more or less reliable balance sheet. Also a change in the number of new shares has to be made public at least in the annual report. Note that the new shares were not issued on the stock market. Instead, the shares were sold to old shareholders. Looking at the trading pattern, in this case, reveals additional information about the shape of the stock market during an event. Obviously, shortly before the public announcement there is a demand surplus that later drives stock prices up. Moreover, the scale of fluctuation is very pronounced during a short time range around the announcement. Also the distance between the demand and supply side of the market widens. That in turn indicates the presence of a high degree of information asymmetry.

Figure 4.12: The case of `Schleswiger Bank´

The reconstructed offer and bid price of the `Schleswiger Bank´ are depicted. Note that I model the demand and supply side of the market based on the information provided by the daily newspaper.



Together with my former results regarding insider-trading, this provides additional evidence about information asymmetry around merger announcements.

4.8 Conclusion

This chapter investigates some extensions of my former event-studies with regard to the impact of announced mergers on stock prices during the first and second phase of globalization. I highlight that some critical points of my event-studies like the change of the estimation period, exogenous time shocks, and the usage of the CMR model are of minor importance. Even after considering the non-synchronous trading problem, the results stay valid. Most notably, the market model applied to the sample of the year 2000 does not yield to different outcomes compared to the CMR model. Note that the market fluctuations observed in the year 2000, as measured by standard deviations of daily returns, are several times larger than during my historical time period. Hence, even in a very volatile market, the stochastic market model leads to similar abnormal returns as the CMR model.

Besides these consistency checks of my event-study results, I also try to explain the time pattern of abnormal returns using the set of explanatory variables that were useful to determine the success of a merger (see chapter two and three). Unfortunately, lacking information regarding the daily trading volume and other relevant daily statistics limits this analysis. One way out is the search for structural breaks in my panel data setting. Once, structural changes are found, a set of dummy variables for the days of the event period as well as for the way of disclosure could be used to explain the adaptation process. But even newly developed methods like the OLS based CUSUM test fail to uncover structural changes in the time dimension. Using a set of n critical regions, one can identify several structural breaks on the individual level. Unfortunately, modeling this breaks with dummies yield to over-specification because abnormal returns of firm i at day t are separately explained with specific dummies. This finding underlines the limited nature of this approach.

Time series analysis, namely the ARIMA model, provides a completely different view on the same data and the some problem. An ARMA specification describes the dynamics of the sequence of abnormal returns and can also be developed further to locate an `empirical event day´. This leads to a modified transfer function approach. Dependent on the way the transfer function is defined, a cross-correlogram identifies the `empirical event day´. This event day is prior to the official newspaper announcement in the case of hidden mergers; therefore, the market anticipates the event respectively private information is conveyed in the order stream. Obviously, there are many pitfalls of ARIMA models and transfer function models – still they shed some light on the stock market development triggered by events without using an event-study.

In the spirit of ARIMA models, I also use a panel based GARCH model to discuss event-induced uncertainty. A GARCH(1,1) process of standardized abnormal returns underlines the importance of time-varying conditional variance that increases in the presence of announcements.

The last section deals with the problem of non-synchronous trading. One might expect that in a historical time period non-synchronous trading leads to fierce biases of event-studies and ARIMA models. Following the outcomes due to Lo and MacKinlay (1990), I argue that trading gaps do hardly influence the variance and covariance structure of daily returns because the daily mean return is close to zero.

Trading patterns provide useful information about the scale of information asymmetry. This last finding could be further developed using models that can estimate an implicit bid-ask spread (see Roll, 1984). Unfortunately, Roll`s (1984) model does not allow time-varying spreads which rules out the possibility that information asymmetry increases around price-sensitive events. In addition, lacking information about the daily trading volumes makes the application of my former methods (see Kling 2002a, b) inappropriate for decomposing spreads. Nevertheless, this can be an interesting topic for future research.

5. The long-term impact of mergers and the role of macroeconomic shocks

5.1 Extended abstract

My aim is to analyze the impact of mergers on company characteristics like share prices, dividends, and the nominal capital over a long time horizon. For that purpose, I develop a panel VAR approach to capture the dynamic responses in share prices and dividends caused by mergers. Besides mergers, macroeconomic shocks play a predominant role in the period from 1870 to 1913. Taking unanticipated changes in inflation and growth rate into consideration extinguishes the role of mergers. Furthermore, my data confirm the emergence of a merger wave around 1900 that drives itself forward. In addition, mergers are more likely if inflation rates increase unexpectedly. Higher inflation rates are accompanied by a phase of real undervaluation of stocks. Hence, the frequency of mergers and the real undervaluation of companies coincide. Most notably, current and past share prices and dividends of a company do not affect the probability to undertake mergers.

5.2 Introduction

5.2.1 The long-term impact of mergers

Thus far, I concentrated on short-term market reactions due to merger announcements. Despite the insights gained by observing daily stock returns, this short-run analysis is limited in the sense that covering a long time period is impossible. Collecting daily data by reading newspapers reaches pretty fast a natural boundary. Hence, to increase the time horizon of my analysis, I draw a sample that consists of annual data on share prices, dividends, and nominal capital. This enables to cover a large time period, namely from 1870 to 1913. In addition, changes in the regulatory environment at the beginning and in the middle of this period make the investigation promising.

Besides the switch from daily to annual data, it is unavoidable – from my point of view – to leave the event-study approach behind. Despite the variety of problems inherent with event-studies when applied to annual data, they are quite common to investigate the long-term effect of mergers. For instance, the event-study of Magenheim and Mueller (1988)[136] tried to shed some light on the permanent influence of mergers. However, using annual data in an event-study approach seems to be precarious because the power of event-studies considerably decreases compared to daily data. Morse (1984) argued that one has to

[136] Magenheim and Mueller (1987) analyzed the abnormal returns of acquiring companies three years after the merger. They found negative abnormal returns and, hence, confirmed the merger paradox.

add hundreds of additional cross-sectional observations to maintain the same power when reducing the frequency of the data. Even if one overcame this technical issue by adding more observations, a causality and omitted variable discussion would turn inference into a risky venture. Observing market reactions that stem from merger announcements within 30 days seems to be reliable. In contrast, arguing that mergers affect the share prices of a company over years without controlling for macroeconomic conditions is quarrelsome. Hence, an event-study approach must be replaced by more sophisticated methods that allow to separate different kinds of shocks on stock prices, dividends, and the nominal capital.

Generally, former studies on mergers during the pre-World-War I period in the USA, Great Britain, and Germany focused on completely different aspects.[137] The interrelation between economic growth, the expansion of large-scale enterprises, and the role of mergers on external growth played a crucial role in these studies. In the spirit of my former studies, the success of mergers is determined by quantifying the market reaction in share prices after the transaction is executed. Furthermore, changes in dividend streams and nominal capital of acquirers can be considered in my long-term study. This enables additional insights whether mergers affect fundamentals like dividend payments and the expansion of enterprises.

5.2.2 Share prices and macroeconomic shocks[138]

To assess the long-run effect of mergers on company characteristics, I develop a panel based VAR framework taking explicitly into account unexpected macroeconomic shocks. For that purpose, it is essential to discuss and quantify the interrelation among macroeconomic shocks, share prices, dividends, and nominal capital.

Whether stocks are an effective inflation hedge, is an old debate in theoretical and empirical investigations and was thoroughly discussed in the 1970s, a period of high inflation. Following the definition of a complete inflation hedge provided by Reilly et al. (1970), the real returns of stock should not be influenced by inflation. In contrast, the inflation illusion hypothesis by Modigliani and Cohn (1979) recently revisited by Ritter and Warr (2002) states that the stock market discounts real dividends at nominal interest rates. Consequently, the market undervalues stocks if inflation is high and overvalues them if inflation is low. Providing empirical evidence whether Reilly's et al. (1970) inflation hedge or Modigliani's

[137] The most prominent studies in this area are Tilly (1982) for the German industry, Davis (1966), Nelson (1959), Eis (1979), Hannah (1974), and Chandler (1977) for the US and UK case. Note that Tilly (1982) did not include mergers within the banking industry, which neglects the predominant role of banks during the pre-WWI period.

[138] See also Baltzer and Kling (2003) for a thorough discussion on the role of macro-shocks.

and Cohn's (1979) inflation illusion hypothesis hold for the period 1870 to 1913 is an interesting contribution to this literature.

Recent studies on the impact of macroeconomic factors on stock markets concentrated mainly on inflation, interest rates and real stock prices. Thereby, one strand of the literature like Tatom (2002) and Rapach (2002) focused on the long-term interrelation between real stock prices and inflation using cointegration analysis. Tatom (2002) found a negative correlation between real share prices and inflation, whereas Rapach (2002) concluded that higher inflation does not erode the long-run real value of stocks. Additional macroeconomic factors were considered that affect real stock prices like interest rates or a diffusion index that represent a set of variables. Besides the long-term effect of macroeconomic variables, studies like Rapach (2001) discussed the short-term reaction of real stock prices caused by changes in macroeconomic conditions. Similar to my approach, he used a structural VAR with a stock price index, money supply, aggregate spending and aggregate supply. In contrast to his long-term study, he found a negative correlation between real share prices and inflation. This finding is typical as mentioned by Anari and Kolari (2001); most studies uncovered a negative impact of inflation on real stock prices in the short-run, whereas in the long-run the outcomes are less clear. The long-run Fisher effect[139] that prices are positively related to inflation was confirmed by Anari and Kolari (2001). They also detected that stock prices react negatively shortly after an increase in inflation – but after some years the effect turns out to be positive. Furthermore, some research was done to detect market reactions triggered by the announcements of macroeconomic news. For instance, Li and Hu (1998) found that new information regarding inflation, employment, and other factors possess considerable impact on share prices after these macroeconomic figures were made public.

Obviously, for the historical time period 1870 to 1913, analyzing announcements is limited caused by data collection problems; thus, I concentrate on the long-term effects of macroeconomic variables on the stock market as discussed by Rapach (2001). Because I cover a relatively long time interval starting 1870 to 1913, an analysis whether share prices have a long-term memory with regard to inflation changes can be carried out. Anari and Kolari (2001) pointed out that share prices possess such a long-run memory.

Accordingly, my first aim is to capture the dynamic responses of share prices, dividends, and nominal capital triggered by unexpected macroeconomic shocks; thereby, as discussed in a subsequent section, I focus on inflation and economic growth rates.

---

[139] Jaffe and Mandelker (1976) discussed the Fisher effect and its implications for stocks in an empirical investigation.

5.2.3 The classical gold standard and macroeconomic stability

My investigation also contributes to the debate with regard to the effectiveness of the classical gold standard[140] to prevent inflation and sustain macroeconomic stability. For instance, Meltzer and Robinson (1988) figured out that the economic uncertainty in the United States was much higher compared to later periods at the turn of the century when the United States was on the gold standard. Uncertainty was higher in inflation rates, nominal GDP, real GDP, and money supply. Results for other countries including Germany[141] are similar, although not as consistent across different economic variables as for the United States.

The gold standard was introduced in Germany in 1873. Besides this major institutional cut, a central bank called `Reichsbank´ was established in 1875. Although, other banks continued to issue currency, the `Reichsbank´ gained more and more control over the money supply in Germany by imposing reserve requirements in foreign exchange and in gold. Henning (1973) argued that these requirements were effective to reduce the possibilities for liquidity creation. Thus, my investigation period covers the beginning of the gold standard and monetary control and ends when the first World War started. However, my focus differs from former studies like Gerhäusser (1990)[142] who concentrated on the impact of inflation rates on the variability of relative prices. In contrast, I try to discuss the dynamic responses of share prices and dividends triggered by unexpected macroeconomic and micro-level shocks.

Why is historical evidence for the interrelation between market responses and unexpected micro- and macro-level shocks interesting for current institutional debates? Since the introduction of the Euro respectively the Asian crisis, there is again a growing attention on the role of exchange rate systems on macroeconomic stability. Moreover, the interrelation among `hot money´ so called portfolio or short-term investments, flight of capital and macroeconomic disturbances is a major concern nowadays. Dibooglu (1998), for instance, discussed macroeconomic stability over a very long time horizon; thereby, his starting point is the period of the classical gold standard. He found that incidence of real demand shocks and monetary shocks are lower under a system of fixed exchange rate.

---

[140] An excellent review article that referred to the NBER conference proceedings entitled "A retrospective on the classical gold standard, 1821-1931" was written by Van Huyck (1987).

[141] Gerhäusser (1990) provided evidence that Germany exhibited remarkable periods of inflation under the classical gold standard lasting from 1873 to 1913. His study is based on Hoffmann's (1965) macroeconomic time series. Furthermore, he concentrated on the interrelation between inflation and the change in relative price variability. Generally, he found that only in periods of deflation variability and with that uncertainty increases.

[142] The data sources and method used by Gerhäusser (1990) were heavily criticized by Borchardt and Rischl (1991).

5.2.4 The emergence of merger waves

When one asks whether my empirical models are based on a theoretical framework that explains the emergence of a merger wave or mergers in general, I have to respond that appropriate theoretical models are still lacking. Generally, theoretical approaches to understand horizontal mergers are seldom.[143] However, the game theoretical model developed by Böckem (2002) is an exception. Unfortunately, there is no way to transfer her heterogeneous Cournot game into an empirical model. The driving force for her horizontal merger wave stems from the heterogeneity in marginal costs among companies. Measuring marginal costs, however, is an unsolvable task, especially in the banking industry, which is mainly responsible for the merger wave.

A recent contribution by Schenk (2001) tried to explain the `merger paradox´ and merger waves using a game theoretical model in which managers apply the minimax-regret decision rule. He argued that a booming economy is the prerequisite to start a merger wave. After the first merger occurred, incentives to imitate this decision increase regardless whether the first merger will be successful. This behavior follows the assertion that one prefers to fail conventionally than to succeed unconventionally. Correspondingly, imitating unsuccessful mergers causes less regret than overlooking the opportunity for a successful merger. Fortunately, an empirical proof based on Palacios-Huerta's (2003) considerations seems to be possible. Detecting a serial dependency of mergers would confirm Schenk's (2001) view. Moreover, whether mergers are successful or not is not essential for his game theoretical model.

5.2.5 Structure of this chapter

The remainder of this chapter is organized as follows. First, I discuss the method of sampling which is based on the paper written by Baltzer and Kling (2003). Thereafter, I construct a panel probit model to anticipate mergers. This enables to evaluate whether a merger is an unexpected shock for the market. Second, I clarify the VAR model used in Baltzer and Kling (2003) – but I extend it to capture changes in nominal capital and microeconomic shocks. Third, I present and discuss my empirical findings followed by a brief conclusion.

---

[143] Böckem (2002) stated that this is mainly due to the lacking profitability of horizontal mergers in a simple Cournot setting with homogenous goods if more than two firms interact. This is also a general finding of the theoretical literature on horizontal mergers. Furthermore, cost asymmetry is necessary to induce profitable mergers in the case of more than two firms (see Salant et al., 1983, Levin, 1990, and Farrell and Shapiro, 1990).

5.3 Method of sampling

5.3.1 How to construct a representative sample?

Generally, there are two methodologies to draw a sample in economic history. In my short-term studies, I chose a limited period of time and collected all available information on mergers during this period. However, in my long-term study, I follow the procedure of Tilly (1982), Weston (1953), and Huerkamp (1979) by selecting companies that fulfil specific criterions like firm size. All of these studies have in common that they focused on surviving companies. Driven by econometric needs, I do the same. However, I will carefully discuss whether the survivorship bias is relevant for my results. Unfortunately, emphasizing the survivor bias was neglected in the former studies mentioned above.

My aim is to construct a sample consisting of 35 leading companies listed on the Berlin stock exchange during the whole period under investigation from the early 1870s to the beginning of the first World War in 1914. By construction of a long-run study, the data set contains only acquiring companies. Restricting the sample to the largest companies ensures that I capture the most active acquirers.[144] A simple approach that includes the 35 largest companies as measured by the paid-in nominal capital would lead to an overrepresentation of banks. Especially, the newly developing industries like the chemical industry would be neglected. To assess mergers and macroeconomic shocks in a variety of industries, I have to construct a sample that is not only limited to the banking industry.

To get an appropriate representation of the German stock market, I divide the listed companies into four major sectors. This procedure is in line with the contemporary division made in `Saling's Börsen-Papiere´ since the early 1870s, namely banks, mining companies, traffic companies, and other industries.[145] I skip the insurance sector because regulations led to a very illiquid trading. This was caused by strict legal requirements concerning the trading in these shares. Changes of ownership must be announced and permitted by the board of directors of the respective company.

To determine weights for every line of business, I have to think about criterions like the number of companies within an industry as used by Tilly (1982). The importance of the different lines of business as measured by the nominal capital decisively changed during the

[144] Tilly (1982) argued that the companies laws of 1884 and the new exchange law established 1896 favored the acquisition of smaller companies by larger companies. This assertion can be justified by the requirement that the minimum issue volume had to exceed one million Mark. Hence, a larger company had advantages to finance acquisitions by issuing new shares.

[145] See `Saling's Börsen-Papiere (1874-1876)´. Strictly speaking, Saling combined the mining sector and the `other industries´ into one category. However, considering the overwhelming importance of the mining industry in the pre-World-War I economy, I prefer to build up an own category for mining companies. Moreover, I refine the `other industries´ into seven sub-sectors: breweries, real estate companies, chemical industry, metal-working industry, mechanical engineering, textile industry and other industries.

44 years. Most notably, the traffic industry and the other-industries-sector underwent a pronounced development. Mainly, because of the nationalization of nearly all railway companies, the contribution of the traffic sector to the total amount of nominal capital decreases rapidly from about 30 % to 10 %. In contrast, I observe a large relative increase of the more and more diversifying other-industries-sector that went along with the proceeding industrialization from around 15 % in 1873 to more than 40 % in 1912. Table 5.1 provides an overview with regard to the change in nominal capital of German companies. But relying solely on this figure does not guarantee an appropriate representation of the German stock market. Especially, infant industries like the chemical industry would be dropped out of my sample if I stick solely to the nominal capital criterion. Inspiring another figure, namely the number of companies,[146] reveals an enormous discrepancy within the four major sectors between the nominal capital and the new measure. Table 5.1 shows that using the total number of companies as criterion, the other-industries-sector played the biggest role in 1873 as well as in 1912. This might be explained by the technological process during this period that enables the emergence of new industries. Obviously, these infant industries started as small companies at the beginning of my investigation period. However, they exhibited a tremendous growth in nominal capital until the year 1912.

Table 5.1: The nominal capital and number of companies in different lines of business

This table presents the paid-up nominal capital and the number of companies at the beginning and the end of the investigation period; thereby, I distinguish among different lines of business.

Nominal capital of German companies in different lines of business			
	1873	1912	Change in percentage points
Banking	44.90%	30.71%	-14.19
Mining	9.43%	16.89%	+7.46
Traffic	30.77%	10.57%	-20.20
Other industries	14.90%	41.83%	+26.93

Total number of listed German companies in different lines of business			
	1873	1912	Change in percentage points
Banking	31.06%	15.03%	- 16.03
Mining	12.97%	5.35%	- 7.62
Traffic	11.26%	10.31%	- 0.95
Other industries	44.71%	69.31%	+ 24.60

[146] This criterion was used by Tilly (1982) to weigh different lines of business.

Besides the technological progress, the legal environment paved the ground for new industries. The new law concerning the foundation of companies passed in 1870 caused a real flood of new companies. During the prosecuting period till 1873, more than 50 % of the newly founded companies belonged to the expanding other-industries-sector. These companies possessed only a low amount of paid-up capital as the new law admitted that the investor had to pay only a fraction of the `official´ nominal capital. Accordingly, this period called `founder boom´ (`Gruenderboom') exhibited an extreme expansion and a large number of foundations.

As my study starts in 1870, I want to capture this wave of newly founded companies. Therefore, I weigh both criterions equally, namely the nominal capital and the absolute number of companies, to decide about the number of companies of each industry that should be included in my sample. In 1873, the turning-point was reached on the stock market accompanied by decreasing founding activities in the following years. Therefore, I fix the year 1873 to collect the data for both criterions from `Saling's Börsenpapiere (1874-1876)´. To take the changes during the whole period into account, I fix as second reference year 1912 close to the end of my investigated period. Data for the year 1912 are provided by the `Kaiserliches Statistisches Amt (1913)´. Consequently, selecting 35 companies by using both mentioned criterions for the years 1873 and 1912 leads to the following numbers of companies for each major sector: 11 banks, four mining companies, five traffic companies, and 15 firms belonging to the other-industries category.

As most of the founding activities took place after 1870, the selection period for my sample also covers companies founded in 1870 or 1871. Note that the criterion for selecting the single companies within a line of business is the nominal capital; thereby, companies with the highest paid-up nominal capital are selected.

Because information about stock specific trading volumes is lacking, I use the companies with the largest nominal capital to select the most important and well-known companies. These `blue chips´ summarized in table 5.2 should represent the most actively traded stocks on the Berlin stock exchange. Consequently, I collect annual share prices, dividends and nominal capital for these `blue chips´ based on `Saling's Börsen-Jahrbuch (1913/1914)´ and the `Handbuch der deutschen Aktiengesellschaften (1911/1912)´.

Table 5.2: Selected companies for my study divided into different lines of business

The date of the last observed share price is set in parentheses. The disappearance of railroad companies is due to nationalizations.

Sector A: Banking | **Replaced companies**

01	Berliner Handelsgesellschaft		
02	Darmstädter Bank für Handel und Industrie		
03	Disconto-Gesellschaft Berlin		
04	Deutsche Bank		
05	Schaaffhausenscher Bankverein		
06	Preussische Bank/Reichsbank		
07	Sächsische Bank		
08	Preussische Bodencredit Actienbank		
09	Allgemeine Deutsche Creditanstalt		
10	Mitteldeutsche Creditbank		
11	Schlesischer Bankverein	36	Deutsche Unionbank (1873)

Sector B: Mining

12	Bochumer Verein für Bergbau und Gussstahl		
13	Laurahütten-Gesellschaft		
14	Phönix		
15	Eschweiler Bergwerksverein		

Sector C: Traffic

16	Berlin-Charlottenburger Strassenbahn	37	Bergisch-Märkische Bahn (1880)
17	Grosse Berliner Pferdeeisenbahn AG	38	Köln-Mindener Bahn (1880)
18	Norddeutscher Lloyd	39	Rheinische Bahn (1880)
19	Allgemeine Berliner Omnibus AG	40	Thüringische Bahn (1880)
20	Aachen Maastricht	41	Hessische Ludwigsbahn (1896)

Sector D: Other Industries

21	Actien-Bauverein „Passage“ (real estate)	42	Dt. Eisenbahn-(Bau-)Ges. (1879)
22	Süddeutsche Immobiliengesellschaft (real estate)		
23	Böhmisches Brauhaus Knoblauch (brewery)	43	Tivoli Brauerei-Ges. (1880)
24	Berliner Unionsbrauerei (brewery)		
25	Ravensberger Spinnerei (textile)	44	Cöpnicker Chem. Fabrik (1874)
26	Schlesische Leinenindustrie-Gesellschaft Kramsta (textile)		
27	Maschinenbau-Gesellschaft Schwartzkopff (machinery)	45	Oberschl.Eisenbahnbedarf(1880)
28	Sächsische Maschinenfabrik Hartmann (machinery)		
29	Ludwig Löwe & Co. (metal)	46	Pollack-Schmidt (1874)
30	Aktiengesellschaft vormals Frister & Rossmann (metal)		
31	Egestorff's Salzwerke (chemical)		
32	Chemische Fabrik Schering (chemical)	47	Schles. Tuchfabrik (1874)
33	Deutsche Continental-Gas-Gesellschaft zu Dessau (others)		
34	Stärkezuckerfabrik Köhlmann (others)	48	Tabacks-Ges. „Union“ (1880)
35	Deutsche Spiegelglas AG (others)		

5.3.2 Inflation rates and economic growth

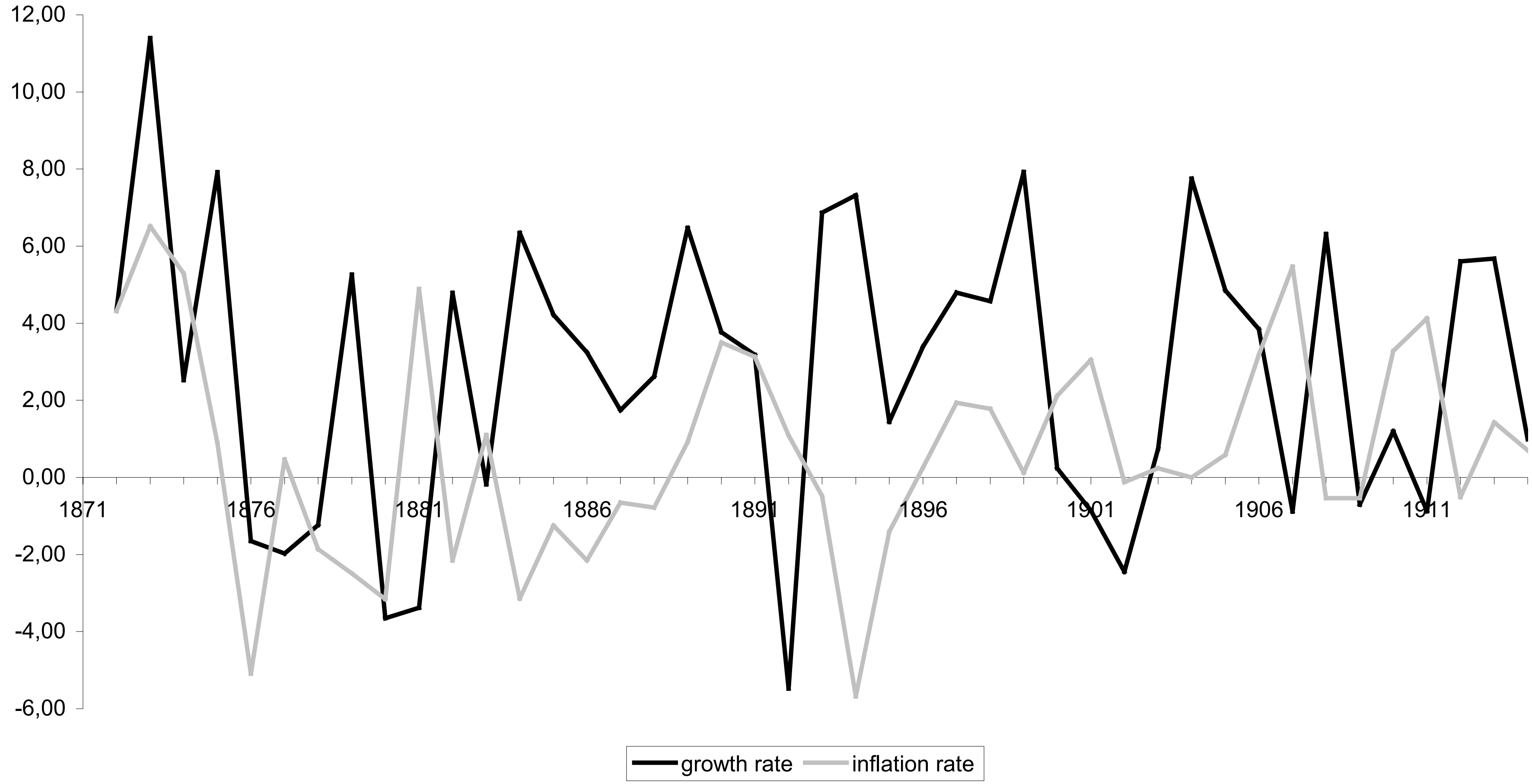
Considering that my investigation is a long-run study embracing a period of 44 years, I have to take the price development into account.[147] Besides the general necessity to deflate long-term time series, my aim is to assess the influence of macroeconomic shocks on share prices and dividends. Hence, I need reliable data on relevant macroeconomic variables; thereby, I focus on inflation and economic growth rates. Possible price deflators are offered by Jacobs and Richter (1935) and by Hoffmann (1965).[148] The Jacobs and Richter index is constructed from wholesale prices, whereas Hoffmann's private consumption index is based on a larger set of time series. Therefore, I decide to use the Hoffmann index – despite the discussible weak points in constructing the data for the 1870s (see Fremdling, 1991, p.41).

My decision is encouraged by Tilly (1992) who also used Hoffmann's private consumption index for deflating his indicator for determining asset returns of the German stock market. Tilly (1992, p.220) concluded that the almost identical course of nominal and real series points to the fact that the development of prices does not seem to dominate the time series at all.[149] A comparison of my real and nominal data supports this view which is not surprising in periods[150] during which the `Mark´ was tied to the gold standard. Prices increased only by 30 % over the whole period. Consequently, the average annual inflation rate was only 0.62 % between 1870 and 1913.

[147] Comparable other long-run studies like Campbell and Shiller (1987) used deflated data as well.
[148] Burhop and Wolff (2002, 2003) tried to correct for some biases of the Hoffmann (1965) time series. However, using the corrected time series leads to quite similar outcomes. My impression is that the differences among alternative price series are more important in levels than in first differences. As my analysis is based on first differences, the results are pretty robust when changing the relevant price index.
[149] Especially studies dealing with a comparison of emerging stock markets give reasons for not deflating their time series by pointing out that the strong devaluation of local currencies to the US dollar would cover all other influences (see Jochum et al., 1999). However, following this procedure is highly disputable.
[150] 1876 to 1913.

Figure 5.1: The real growth rate of net national product and inflation rates 1870 to 1913

The investigation period exhibits a tremendous scale of macroeconomic fluctuation; thereby, the data are due to Hoffmann (1965).



Nevertheless, inflation rates exhibit a tremendous fluctuation during the period. Inspiring figure 5.1 underlines that remarkable periods of deflation and inflation existed that could severely affect stock prices and dividends.[151] Besides the inflation rates, figure 5.1 also depicts the growth rate of net national product in real terms. Both series underline the remarkable fluctuation in macroeconomic conditions during this period.

5.3.3 Testing for unit-roots in share prices and dividends before and after deflating

Deflating share prices and dividend series affects incisively the time series characteristics. Note that I obtain ambiguous results applying individual unit-root test to nominal share prices and dividends. Nominal dividends are typically I(0) processes,[152] whereas after deflating both series, real share prices and dividends, are predominantly I(1)-processes. Considering the relative weak power of individual unit root tests based on 43 observations, I want to conduct panel unit root tests as discussed later. For that purpose, I have to close sporadic gaps in my data set as the panel unit root tests demand complete time series without any gaps.

5.3.4 Missing values and the Holt-Winter filter

As a common problem of empirical studies, I have to deal with gaps in my individual time series. Apart from some `naturally´ missing values in the first years because of not-yet-listed companies, nine share prices are missing between the founding year of the company and the end of my investigation period. Therefore, the missing value rate is clearly below 1%. As the missing values are not concentrated on one or two single companies, they neither cause any distortions of my VAR-analysis nor of tests concerning single time-series. As mentioned above, for panel-based unit root tests, I need a sample without any gaps. Hence, I decide to use the Holt-Winters (see Winter, 1960) exponential non-seasonal smoothing method to fill the gaps. The exponential smoothing seems to fit very well to my annual share price series as this method is appropriate for series with a linear time trend and no seasonal variation. Nevertheless, other procedures like the Hodrick-Prescott filter or the calculation of the average based on the previous and subsequent observation of the missing value lead to similar outcomes.

---

[151] The longest deflationary phase started in the middle of the 1870s and lasted for almost 10 years with a price decrease of up to 5% in 1875.

[152] ADF and KPSS point in opposite directions; hence, the results are said to be uninformative for the time series of dividends.

5.3.5 Annual data on mergers: Discussing the pros and cons

Of course, working with annual data and the `Handbuch der deutschen Aktiengesellschaften´ as information source for executed mergers is crude compared to my former short-term analysis. A long-term study makes it impossible to assess the effect of mergers on target firms. Furthermore, unsuccessful mergers announced in the daily newspapers – but not executed later cannot be considered in my long-term study. However, the problem of a `spotlight´ analysis which focuses on a specific year can now be solved. But this solution comes at a high cost, namely the survivorship bias and the need to control for macroeconomic fluctuations. Nevertheless, both issues can be thoroughly discussed and sufficiently solved. Consequently, the short-term as well as the long-term view can both contribute to increase the understanding of mergers by highlighting different aspects. To give a first impression, table 5.3 summarizes the executed mergers; thereby, the 35 companies are in all cases the acquirers. In contrast to the last merger wave which took place shortly before the `new economy bubble´ burst in the year 2000, mergers among large companies were not common in the pre-WWI period. At a first glance, the banking industry accounts for the overwhelming part of all mergers. This finding is in line with my short-term study.

Table 5.3: Mergers executed by the respective acquirer, 1870 to 1913

Smaller transactions like the purchase of a single branch or a production facility in a specific city are skipped. In addition, collusive arrangements like pooling agreements (`Interessen-gemeinschaften´) that do not lead to a full merger in the legal sense are excluded.

Sector A: Banking	
Berliner Handelsgesellschaft	Internat. Bank (1891)
Darmstädter Bank für Handel und Industrie	R. Haussig (1900), H. Oppenheimer und O. Davisson (1901), Bank für Süddeutschland (1902), Breslauer Disconto Bank (1902), Robert Warschauer & Co (1904), Hermann Arnhold & Co (1906), Ed. Loeb & Co (1907), Commandite Wingenroth, Soherr & Co (1909), America-bank AG (1909), Bayerische Bank für Handel und Industrie (1910), J. Sander (1910), Kohrs & Seeba (1911)
Disconto-Gesellschaft Berlin	Norddeutsche Bank (1895), J. Schultze & Wolde (1904), Schlieper & Co (1906), Gebr. Neustadt (1907), Meyer Cohn (1908), Bamberger & Co (1909), L. Mende (1911)
Deutsche Bank	Frankfurter Bankverein (1886), Menz, Blochmann & Co (1901), Bühler und Heymann (1906), Balser & Co (1910)
Schaaffhausenscher Bankverein	A. & L. Camphausen (1903), Niederrhein. Kredit-Anstalt, former name: Peters & Co (1904), Westdeutsche Bank (1904)
Preussische Bank/Reichsbank	-
Sächsische Bank	-
Preussische Bodencredit Actienbank	-
Allgemeine Deutsche Creditanstalt	Becker & Co (1901), Günther und Rudolph (1903), Kunath & Nieritz (1905), Vereinsbank zu Grimma (1905), Bernburger Bankverein (1907), additional smaller acquisitions in 1908/1909
Mitteldeutsche Creditbank	B. Berlé (1898), Aron Heichelheim (1906), Arthur Andrea & Co (1906), Moritz Heertz (1906), Herm. Wertheim (1906), North Kammeier & Co (1908), Gebr. Fürth & Co (1909), Bernard Weinmann (1910)
Schlesischer Bankverein	Abraham Schlesinger (1905)
Sector B: Mining	
Bochumer Verein für Bergbau und Gussstahl	-
Laurahütten-Gesellschaft	Ges. Eintrachthütte (1894), Siemanowitz, Baingow and Przelaika (1904)
Phönix	Westphälische Union zu Hamm (1898), Hoerder Verein (1907), Akt. Ges. Steinkohlenbergwerk Nordstern (1907), Düsseldorfer Röhren- und Eisenwerke (1910)

Eschweiler Bergwerksverein	Vereinigungs-ges. für Steinkohlenbau im Wurmrevier (1907), Eschweiler-Köln Eisenwerke AG (1910)
Sector C: Traffic	
Berlin-Charlottenburger Strassenbahn	Grosse Berliner Strassenbahn (1900)
Grosse Berliner Pferdeeisenbahn AG	Neue Berliner Pferdebahn (1900)
Norddeutscher Lloyd	-
Allgemeine Berliner Omnibus AG	Neue Berl. Omnibus-Gesellschaft (1903), Victoria-Speicher AG (1905)
Aachen Maastricht	-
Sector D: Other Industries	
Actien-Bauverein „Passage“ (real estate)	-
Süddeutsche Immobiliengesellschaft (real estate)	-
Böhmisches Brauhaus Knoblauch (brewery)	-
Berliner Unionsbrauerei (brewery)	Eberswalder Aktienbrauerei (1906), Klosterbrauerei Charlottenburg (1909)
Ravensberger Spinnerei (textile)	-
Schlesische Leinenindustrie-Gesellschaft Kramsta (textile)	-
Maschinenbau-Gesellschaft Schwartzkopff (machine)	-
Sächsische Maschinenfabrik Hartmann (machine)	
Ludwig Löwe & Co. (metal)	-
Aktiengesellschaft vormals Frister & Rossmann (metal)	-
Egestorff's Salzwerke (chemical)	Kiesbaggerei Rohrsen-Drakenburg (1896), Nieburger Fabrik (1909)
Chemische Fabrik Schering (chemical)	-
Deutsche Continental-Gas-Gesellschaft zu Dessau (others)	-
Stärkezuckerfabrik Köhlmann (others)	Factory in Schneidemühl (1880), factory in Fürstenwalde (1882)
Deutsche Spiegelglas AG (others)	-

5.3.6 Share prices, dividends, and nominal capital in different industries

To illustrate the development of real share prices and dividends in different industries, figure 5.2 and 5.3 plot the respective time series. Generally, one recognizes a high degree of co-movement in share prices, whereas the mining industry exhibited a conspicuous fluctuations in dividend payments. The high variability of dividends stresses the high dependency of dividend payments on current earnings. However, reported earnings and actual earnings can

Figure 5.2: Development of the real share prices 1870-1913 in different industries

I plot the average share price of the respective industry; thereby, all values are expressed in prices of the year 1913. To construct the index values, the year 1870 is chosen as reference basis.

Figure 5.3: Development of the real dividends 1870-1913 in different industries

I plot the average dividends of the respective industry; thereby, all values are expressed in prices of the year 1913. To construct the index values, the year 1870 is chosen as reference basis.

deviate considerably in the pre-WWI period. Nevertheless, collecting dividends is easier and due to missing accounting standards more reliable than determining earning per share.

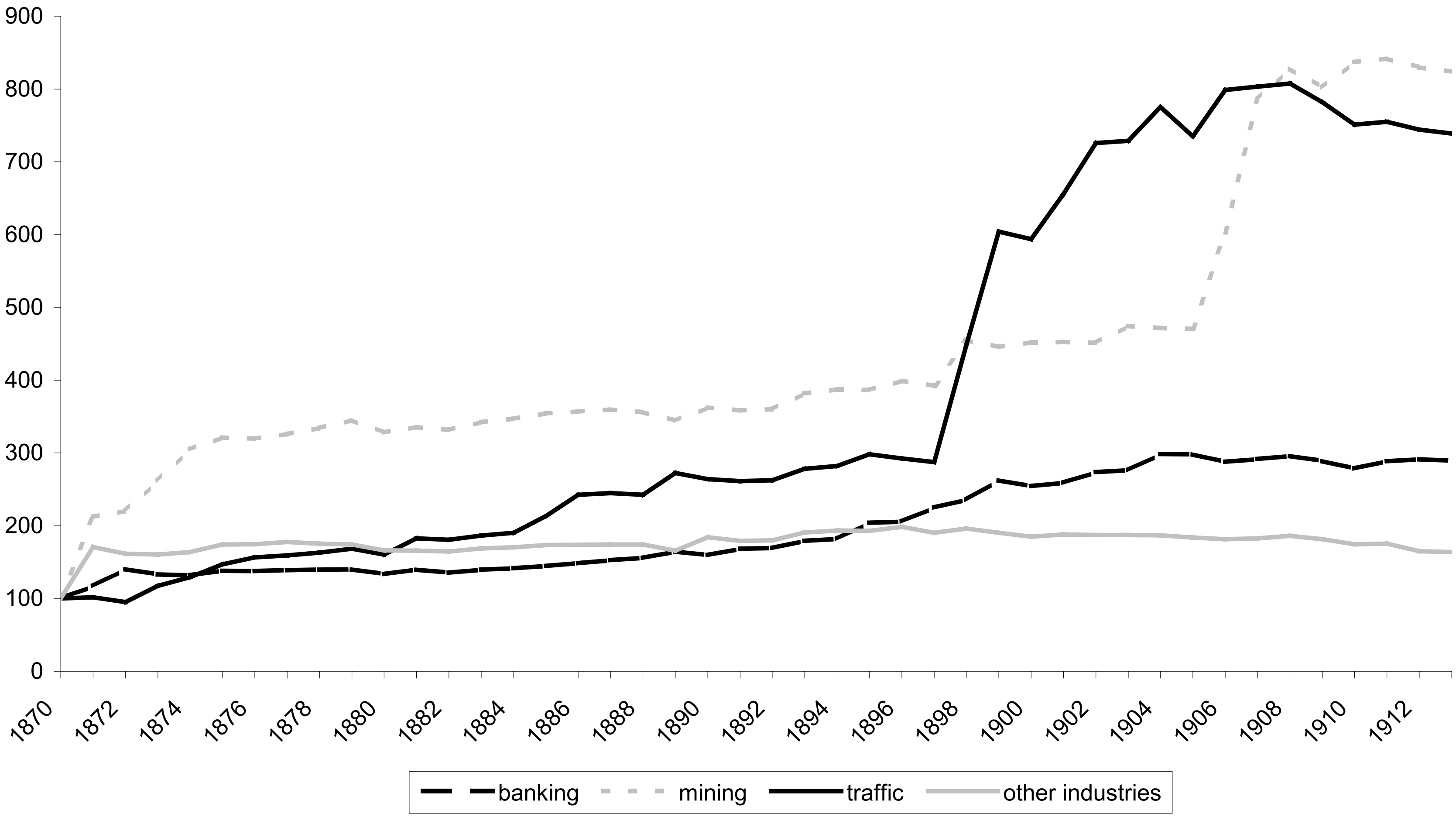
Obviously, the period 1870 to 1913 showed a steady expansion of enterprises in terms of real nominal capital. Figure 5.4 depicts the development of real nominal capital in different lines of business. Although banks were very active in initiating mergers, the development of nominal capital is moderate compared to other industries. As depicted by Figure 5.4, there was a considerable discrepancy in the development of the real nominal capital of different lines of business from 1870 to 1913. While some industries like the mining sector grew rapidly, other industries did not show an upward tendency. Most noteworthy, banks were on average 8.28 times larger than the standard company belonging to another industry. This size ratio mitigated over time. In the year 1913, the ratio declined to 6.20; hence, other industries exhibited a catch-up growth during this period. An upsurge in the nominal capital can stem from internal growth or from external growth through mergers and acquisitions. To what extent mergers are responsible for the expansion path, will be discussed thoroughly in a later section.

5.3.7 How important is the survivorship bias?

As I want to discuss the long-term fluctuations in the German capital market affected by mergers and macroeconomic shocks, one should stick to the initial cross-sectional units. As a panel VAR approach is used later, all companies should survive during the investigation period to guarantee that I can observe share prices, dividends and nominal capital without gaps. Note that missing values may lead to a reduction of the optimal lag length of the VAR. Hence, the possibility to obtain reliable estimates for the long-term dynamic would be limited. Obviously, one practical solution to tackle the survivorship bias is to construct portfolios for every line of business. The composition of these portfolios may vary over time; hence, newly founded companies and bankruptcies are taken into account. From my point of view, building

Figure 5.4: Development of the real nominal capital 1870-1913 in different industries

I plot the average nominal capital of the respective industry; thereby, all values are expressed in prices of the year 1913. To construct the index values, the year 1870 is chosen as reference basis.



portfolios has three major shortcomings that rule this possibility out for my analysis. First, one has to deal with a causality problem that would arise if the companies were switched during the period. Detecting a severe change in share prices could have two reasons. A change could stem not only from an exogenous shock like an increase in inflation but also from a change of the composition of a portfolio. So it seems reasonable to avoid this problem. Second, when using, for instance, four portfolios for the major sectors, one does not exploit the full information contained in the panel data set. Even worse, the optimal lag length of the VAR model would not exceed one respectively two.[153] Generally, the more individual time series are included, the more complicated VAR model can be estimated, and the more precise are the results for the long-term dynamics. Finally, evaluating the impact of mergers – a micro-level shock – should be done on the company level and not for the whole industry. Consequently, there are good reasons to capture the dynamics on the company level – but having in mind the inherent survivorship bias.

Nevertheless, by neglecting companies that leave the market, I am aware of the fact that a survival bias could affect my results. To deal with this issue, I construct an unbalanced panel in which companies are included that fulfil the same criterions mentioned in section 5.3.1 without the presupposition that the chosen company has to be listed for the whole period. Following this procedure, it turns out that twelve additional companies have to be considered. Table 5.2 also reports these companies. Within the transportation sector, the whole sample has to be replaced because all initially selected companies in the year 1870 went into bankruptcy or were nationalized. Until their nationalization at the end of the 1870s, the transportation sector was dominated by the big railroad companies. Most of the replaced companies of the other-industries-sector either failed or were delisted within the first decade of the investigated time period.

How could the survivorship bias affect my results? By looking at surviving companies, the severity of micro- and macro-level shocks might be understated. Fortunately, I can make an assessment of this effect. By using a probit model to test the impact of the same variables used later in my VAR models on the possibility of a company to fail, I do not get any significant result.[154] Therefore, the survivorship bias seems to be less severe. However, most of the non-surviving companies were delisted during the founder crisis; hence, I have only a few observations. This fact limits the possibility to detect the influence of macroeconomic shocks on the probability to leave the market.

[153] This depends on the selected information criterion.
[154] Results are available from the authors on request.

5.4 What drives merger during the first phase of globalization in Germany?

5.4.1 How should I model the driving forces for mergers?

Before conducting a VAR analysis to measure the impact of mergers on share prices, dividends, and the nominal capital, it is essential to model the decision to merge. Note that mergers cannot be regarded as exogenous microshocks – but are the result of a decision process within a company and their shareholders. Company specific characteristics could play a crucial role for encouraging or preventing future expansion plans. One may argue that inserting the merger decision as additional endogenous variable into a broader VAR framework would work. However, as far as I know, VAR models that allow a binary choice variable as endogenous variable are not yet developed.[155] Thus, I propose the following two-step approach: First, applying a panel probit model uncovers how the merger decision depends on the variables used in the VAR model. Second, I make a prediction with regard to mergers in the following year based on today's knowledge. This enables to identify two types of errors, namely unanticipated mergers and falsely forecasted ones. Note that I mean with the term \`falsely forecasted´ mergers that a merger is expected by market participants – but in fact is not executed. Both errors can be handled as exogenous shocks in the VAR framework; therefore, I eliminate the endogenous nature of mergers.

5.4.2 Merger activity during the investigation period 1870 to 1913

Only a few studies on mergers during this time period are available for Germany. Moreover, as mentioned in section 5.2.1 the interest rests on the role of external growth and not on the success of mergers. Among these studies, Tilly's (1982) contribution is the most noteworthy because he covered the period 1880 to 1913 and focused on large-scale enterprises. Thus, my sample should be comparable to Tilly's (1982) data set. Note that in both cases annual data are used; however, Tilly (1982) worked with a different methodology to choose the cross-sectional units. Consequently, he claimed that only surviving companies were considered - but the number of companies in his sample increased over time. It stays unclear for the reader whether this varying number of observations stems from the inclusion of newly founded firms or is due to lacking information on the expansion of the respective enterprises.

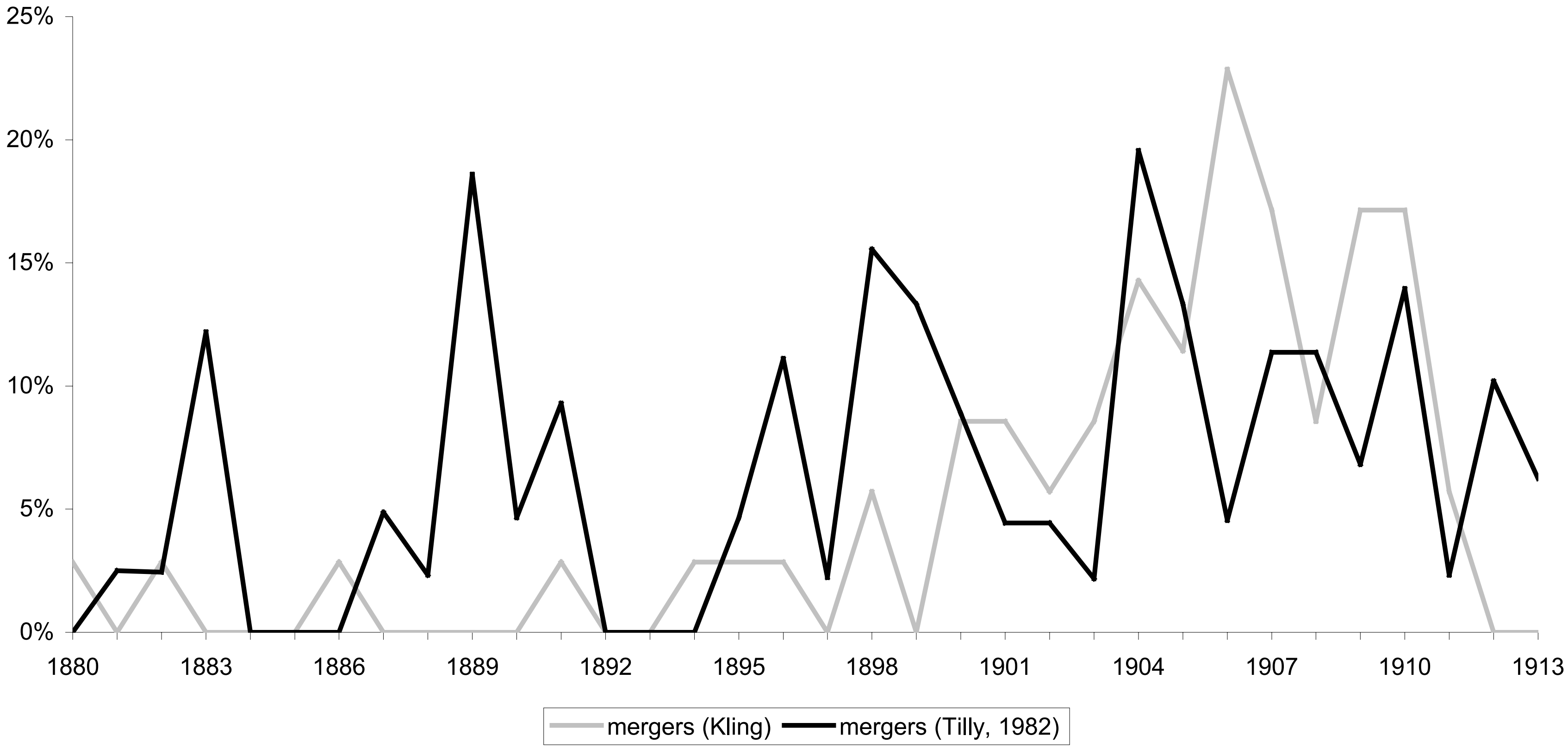
Figure 5.5 provides an overview with regard to the number of executed mergers in my sample. This graph also embeds the number of takeovers provided by Tilly (see p.634, table 1, 1982). Compared to Tilly (1982), my sample shows that the mergers are centered around a peak in the year 1906. Note that Tilly (1982) did not include the banking industry. Moreover,

[155] Note that it is possible to use the standard linear regression model to explain the merger decision – but it is well known that this is misleading for binary variables (see Greene (2000)).

he covered only the period from 1880 to 1913, and his sample size varied from 38 to 49. To take the changing sample size into account, figure 5.5 reports the number of mergers divided by the sample size. This can be regarded as a measure for the merger activity in the sample. The difference could stem from the fact that Tilly (1982) excluded the banking industry. The exclusion matters because in my sample, banks account for 70% of all mergers. In addition, my short-term study emphasized the importance and the success of mergers among banks. Hence, neglecting the banking industry, Tilly (1982) can only tell a part of the story.

Figure 5.5: Merger activity in the period 1880 to 1913

To compare my sample with the results of Tilly (1982), I calculate the number of mergers divided by the sample size. Thus, this is a measure for the merger intensity.



Noteworthy, Huerkamp's (1979) figure about the merger activity is very close to my finding even though she excluded the banking industry. She uncovered a strong increase in the number of mergers from 1887 to 1907 – but her analysis is limited to a pure descriptive study because additional firm specific information was not collected. In contrast to Tilly (1982) and my study, she also included firms that are not listed on stock exchanges.

5.4.3 Model selection of a dynamic panel probit model with random effects

The decision to undertake a merger within one year can be regarded as a count data model if more than one merger is conducted. However, acquirers undertook rarely more than one merger within one year. Hence, it seems to be appropriate to model the binary decision of a company: to merge or not to merge. Because I deal with panel data, a panel discrete choice model should be applied. Fortunately, in recent years considerable progress in estimating panel probit respectively logit models was made.[156] To account for company specific effects, I propose a random effects model of the following shape. Note that this specification can be justified later by running log-likelihood ratio tests.[157]

$$m_{it} = \alpha + \sum_{j=0}^{p} \boldsymbol{\beta}_{\mathbf{j}}' \boldsymbol{\Delta z}_{\mathbf{i(t-j)}} + \sum_{j=1}^{p} \gamma_j m_{i(t-j)} + u_i + \varepsilon_{it} \tag{5.1}$$

Obviously, m_{it} takes the value one if company i executes a merger in year t and zero otherwise. The column vector $\Delta z_{i(t-j)}$ consists of the first differences in real share prices, real dividends, and real nominal capital. Note that I take the natural logarithm before calculating the respective first difference.[158] To determine the lag length of this dynamic model, I carry out the maximum likelihood estimation of (5.1) with different lag specifications. Thereafter, the Akaike and the Schwarz criterions are calculated together with the log likelihood of the respective model. To make a comparison reliable, the relevant sample is fixed.

Based on McFadden (1974), one can derive a pseudo R^2; thereby, a probit model that includes only a constant term have to be estimated to obtain a reference basis for the log likelihood. Hence, the log-likelihood of every model specification j is compared to the log-likelihood of the reference model with constant term denoted ll_0.

$$R^2_{PSEUDO} = 1 - \frac{ll_j}{ll_0} \tag{5.2}$$

156 Hsiao (1992) provided an excellent overview of panel probit models.
157 Note that my distributed lag model could be affected by multicollinearity among lagged explanatory variables – but I concentrate on predicting mergers. Correspondingly, a high explanatory power is more important than precisely determined partial impacts.
158 Note that I take always natural logarithms when I refer to share prices, dividends, or nominal capital.

Table 5.4 reports the outcome of model (5.1) with different lag length p. The Akaike criterion indicates that the lag length should be set equal to two, whereas the Schwarz criterion favors the reference model with constant term. Regardless which lag structure is used, random effects matter indicated by log-likelihood ratio tests. The null hypothesis of no random effects can be rejected in all cases. This models point out that current and former changes in share prices and dividends as well as the growth rate of the net national product denoted gdp_{it} possess no impact on mergers. In contrast, higher inflation rates and mergers executed one or two years ago lead to a higher probability that during the year t company i announce a merger.

Using specification tests like log-likelihood ratio tests (LR),[159] the number of explanatory variables can be further reduced. Finally, the probit model has the following structure.

$$m_{it} = \alpha + \beta \Delta n_{it} + \sum_{j=1}^{2} \gamma_j m_{it-j} + \delta \inf_{it} + u_i + \varepsilon_{it} = \tag{5.3}$$
$$-2.2808 + 0.9726 \Delta n_{it} + 0.8708 m_{it-1} + 0.9366 m_{it-2} + 0.9726 \inf_{it} + u_i + \varepsilon_{it}$$

All coefficients are significant on the 1% level of significance. How can one interpret this result? To facilitate the interpretation, one can calculate marginal effects. A merger in the prosecuting period increases the probability for an additional expansion by 8.08 percentage points. Mergers occurring two periods ago still influence the current merger activity with a marginal effects of about 9.20. Hence, there is evidence for the emergence of merger waves during the investigation period. Furthermore, an increase in inflation rates by one standard deviation adds 1.51 percentage point to the probability of mergers. To obtain an impression whether these marginal effects are essential, one should have in mind that the average forecasted probability for merger reaches 2.93%. Most notably, the current and past changes in company characteristics like share prices and dividends do not matter.

[159] The test statistic reaches 8.58 and the corresponding p-value 0.804.

Table 5.4: Selecting the appropriate dynamic panel probit model

Applying a maximum likelihood estimation procedure, I run random effects probit models with different lags. To calculate the pseudo R^2 statistic, a reference model with a constant term is estimated. The number of stars indicate significance. One star represents 10%, two 5% and three 1% level of significance. The LR statistic tests whether random effects should be assumed; thereby, the null hypothesis states that individual effects can be neglected.

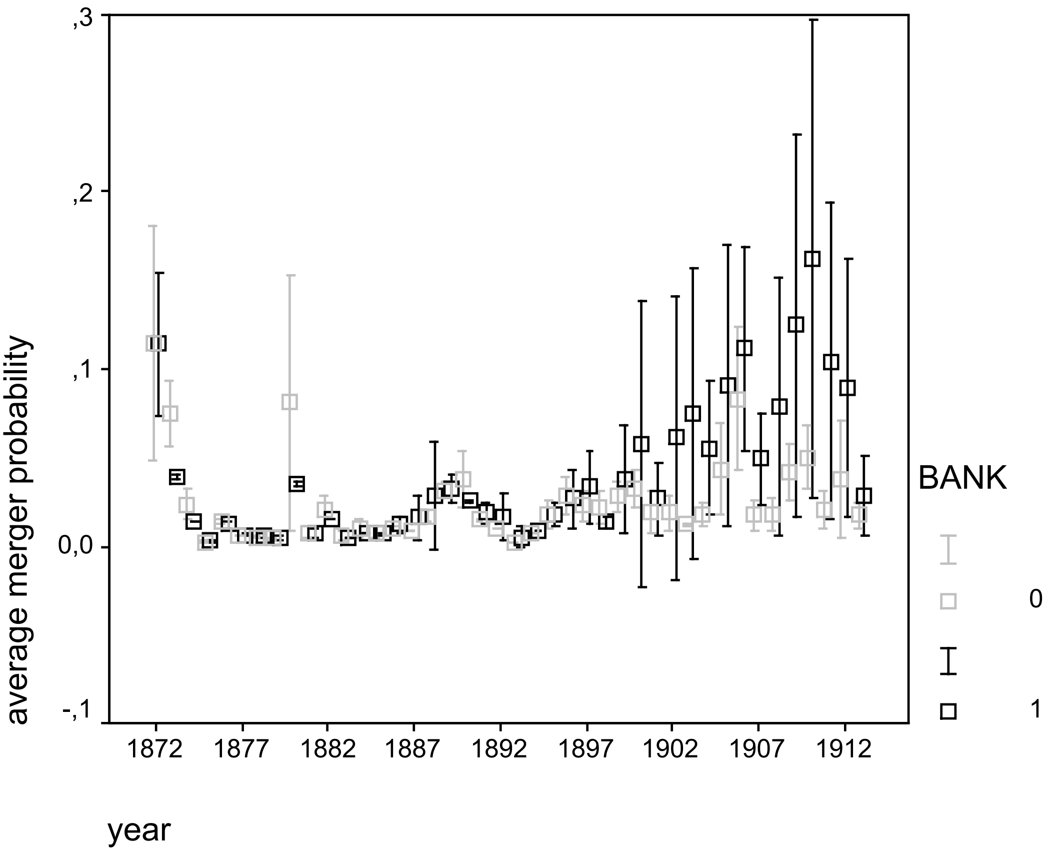
	Basic model	Without lags	One lag	Two lags	Three lags
Constant	-2.0584***	-2.2969***	-2.5611***	-2.5159***	-2.5537***
m_{it-1}	-	-	0.8363***	0.7361***	0.6544***
m_{it-2}	-	-	-	0.8133***	0.7060***
m_{it-3}	-	-	-	-	0.2491
Δp_{it}	-	0.2763	0.8192	0.7474	0.7536
Δp_{it-1}	-	-	0.4246	0.7821	0.8031
Δp_{it-2}	-	-	-	-0.2076	0.0205
Δp_{it-3}	-	-	-	-	-0.0096
Δd_{it}	-	-0.0548	-0.1156	-0.1218	-0.1544
Δd_{it-1}	-	-	-0.3365	-0.3243	-0.3669
Δd_{it-2}	-	-	-	-0.2160	-0.2714
Δd_{it-3}	-	-	-	-	-0.2365
Δn_{it}	-	1.0266***	1.2653***	1.1530***	1.2119***
Δn_{it-1}	-	-	-0.2729	-0.0362	-0.0297
Δn_{it-2}	-	-	-	0.4997	0.6487
Δn_{it-3}	-	-	-	-	-0.4090
$inflation_{it}$	-	0.1895***	0.2064***	0.1871***	0.1874***
$inflation_{it-1}$	-	-	0.0897	0.1160*	0.1364**
$inflation_{it-2}$	-	-	-	0.0336	0.0522
$inflation_{it-3}$	-	-	-	-	0.0008
gdp_{it}	-	0.0027	-0.0117	-0.0147	-0.0217
gdp_{it-1}	-	-	-0.0007	-0.0061	-0.0039
gdp_{it-2}	-	-	-	0.0011	0.0006
gdp_{it-3}	-	-	-	-	-0.0123
AIC	416.95	397.02	385.44	380.20	388.77
SBIC	422.14	428.19	452.98	478.90	518.65
Log-likelihood	-207.47	-192.51	-179.72	-171.10	-169.39
Observations	1333	1333	1333	1333	1333
Pseudo R^2	-	0.07	0.13	0.18	0.18
LR-test	43.42***	48.89***	22.45***	7.81***	6.40***

5.4.4 Predicting mergers during the period 1870 to 1913

In line with my descriptive finding that the merger activity increases after 1895, the probit model predicts higher probabilities for this period. However, the peak in the year 1872 seems to be misleading. Distinguishing between banks, which exhibited the highest activity in these years, and all other industries highlights the remarkable discrepancy between banks and other industries after 1900. Figure 5.6 plots the average expected probability for mergers based on the knowledge available in the previous year. Hence, the figure shows the one year forecast.

Figure 5.6: Predicted probability to merge for banks and other industries

I calculated the predicted probability that a firm will initiate a merger in year t based on the available information at t-1. This forecasts are obtained for every observation in the panel data set. Thereafter, I derive the average predicted probability and plot the 95% confidence intervals for this estimate.



To determine the expected merger, one has to specify a cutoff rate that has to be exceeded to expect a merger in the following year. Using a cutoff value around 0.125 enables to reach the highest possible accuracy for correctly anticipated mergers. However, only 32.08% of all mergers are correctly anticipated by the model; hence, many merger are not predictable. Restricting the relevant time span to the period after 1896, which was characterized by the new exchange law established in the year 1896, yields better forecasts. Note that the same cutoff value is optimal for the reduced time span; however, the probability of correctly anticipated mergers reaches 40.48%.

Consequently, I can now determine which merger can be anticipated and which one act as a microeconomic shock. Merger shocks can be treated as exogenous events that trigger responses in share prices, dividends, and nominal capital. Thus, by inserting the unpredicted mergers into a VAR framework, impulse responses can be obtained.

Obviously, I have to deal with the problem how to handle incorrectly predicted mergers. When I believe in the discrete choice model, then false predictions should lead to disappointments. Therefore, a negative microeconomic shock can affect prices, dividends, and the nominal capital. By defining two types of shocks, one obtains two types of potential errors that can be inserted into the VAR. The dummy variable denoted m_p takes the value one if a merger occurs surprisingly, whereas the dummy m_n stands for false forecasts. Embedding both surprising events into a VAR framework enables to test whether these mistakes trigger any consequences.

5.5 The anticipation of macroeconomic variables

5.5.1 Why is it essential to talk about anticipation of macroeconomic conditions?

Inserting macroeconomic shocks in my VAR framework as exogenous variables is only permitted if it is indeed an unexpected shock after observing the realizations of the endogenous time series of the VAR. Putting this in other words, it states that exploiting the information reflected in past share prices, dividends, and nominal capital does not help to predict macroeconomic shocks. There are two different methodologies to deal with this problem. First and most common, inflation rates and economic growth rates are included besides share prices, dividends, and nominal capital as additional endogenous time series into the VAR. Second, a two step approach is applied; thereby, the unexpected component of macroeconomic variables is determined in a separate model. Thereafter, the unpredicted components are included as exogenous shock into the VAR model. Caused by my model regarding the merger decision, which is by nature also a two step setting, I prefer the latter opportunity to maintain consistency. However, both approaches are technically similar – given that one imposes appropriate ordering restrictions.

5.5.2 Can one anticipate future inflation and growth rates?

Assuming that market participants observe past changes in stock prices, dividends, growth rates, inflation rates and nominal interest rates, one can infer the expected inflation and growth rate. Therefore, I specify an ARIMA model based on the Akaike and Schwarz criterion for inflation and growth rates. The optimal ARIMA specification is for both macroeconomic time series a first order autoregression. Furthermore, the number of lags reaches one for the set of explanatory variables.[160] Inserting lagged values of the changes in the share prices and dividend index as well as information about the nominal interest rates,[161] one obtains an expected value for inflation and growth rate.[162] In line with literature discussing the classical gold standard, inflation and economic growth are hardly predictable.[163] Figure 5.7 plots the observed macroeconomic figures and the unexpected part resulting from a one-step forecast. Apparently, the graphs indicate that almost the whole inflation and growth rate came as a surprise. These unexpected components are inserted as macroeconomic shocks in my VAR model.

[160] Note that this model is close to my transfer function analysis; hence, it is also possible to check the quality of the chosen lags by inspiring cross-correlograms between the respective macroeconomic variables and the explanatory variables.

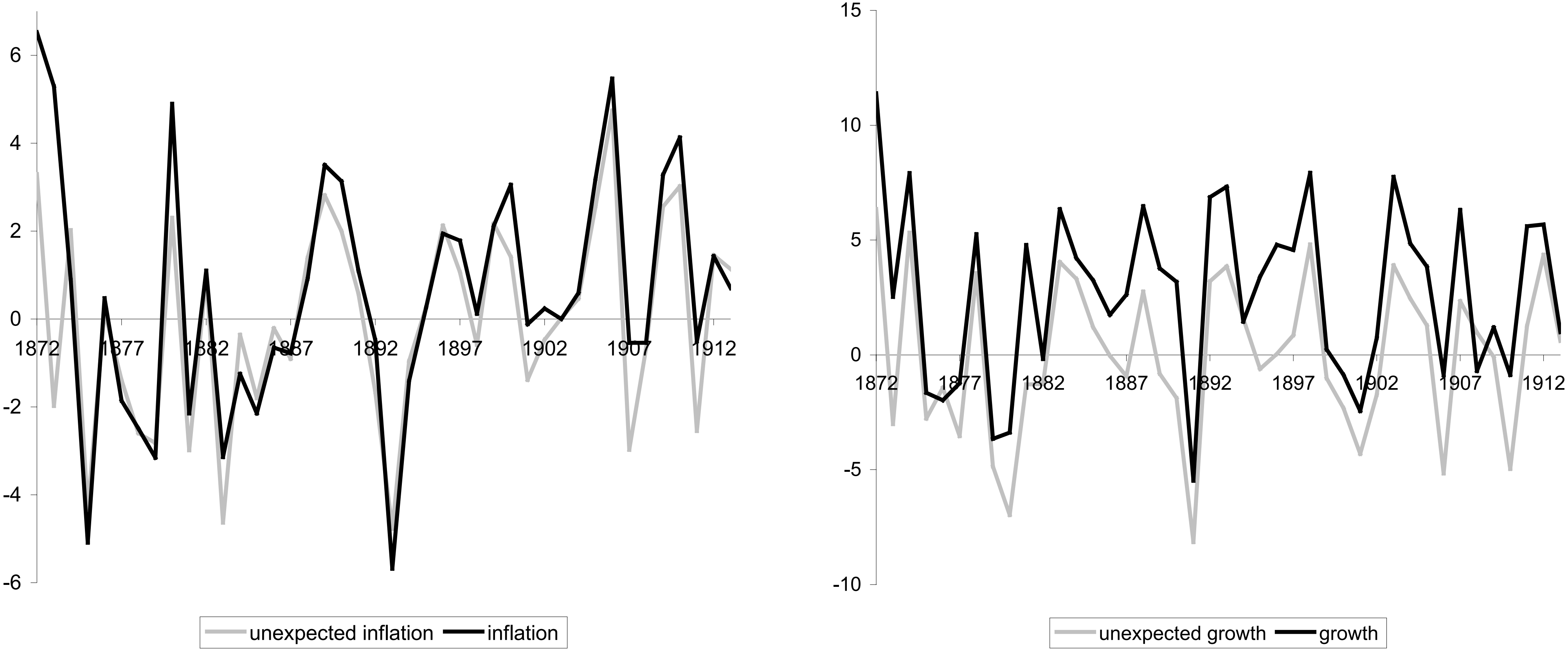
[161] I take the discount rate for private banks; however, all sequences of nominal interest rates show a quite similar movement over time. The results are valid if I use bond yields instead of discount rates. \`Neumann's Kurstabellen´ provide an excellent data source on a monthly basis.

[162] The exact results are available from the author on request. I skip it here because they are of minor importance.

[163] See for a qualitative view on that issue Davis, Neal, and White (2002).

Figure 5.7: Observed and unexpected inflation and growth rates

In panel A, observed and unexpected inflation rates are plotted, whereas panel B shows the result for growth rates of the net national product.



Panel A

Panel B

5.6 Panel vector autoregression with macro and microshocks

5.6.1 `Traditional´ VAR model with prices, dividends, and macroeconomic factors

To capture the interrelations among share prices, dividends, growth rate of net national product, and inflation, I build up a VAR framework. Thereby, I use an index of logarithmically transformed equally weighted share prices as well as dividends. Unfortunately, having only 42 observations – but depending on the optimal lag length lots of parameters – this model is very limited. Using equally weighted indices of share prices and dividends as endogenous series can be seen as the `traditional´ approach to analyze the dynamics triggered by macroeconomic shocks. Unfortunately, this `traditional´ view is limited in at least two ways. First, the number of lags is very low; hence, the influence of macroeconomic shocks cannot be observed over longer horizons. Following the Hannan-Quin, Schwarz, and the Akaike criterion, I specify a VAR in reduced form with one lag. Even this simple lag structure requires to estimate many parameters. As I have only a few observations, the estimates are very inaccurate. Moreover, the bootstrapping intervals are very large for the impulse response functions. Besides the imprecise estimated dynamic responses, mergers should be investigated on the company level. Hence, using equally weighted indices is inappropriate to quantify the influence of mergers.

To overcome these problems, I propose a panel VAR framework to exploit the largest possible amount of information provided by my data set. This enables to analyze long-term effects of macro- and micro-shocks by estimating a further extended lag structure.

5.6.2 Panel VAR framework with share prices, dividends, and exogenous shocks

I try to capture the dynamics between share prices and dividends without taking into account changes in the nominal capital. Section 5.6.4 discusses thoroughly, why nominal capital can be neglected.

Obviously, a VAR model imposes the requirement that the first differences of dividends and share prices must be stationary. This assumption can be confirmed by unit-root tests applied to 70 individual series of share prices and dividends.[164] However, caused by the increase in the number of observations, it seems to be worthwhile to prove these assumptions on the panel data level. In applied empirical research,[165] two different kinds of tests are used; thereby, one category has as null hypothesis that all series are stationary, whereas the other category uses non-stationarity of all series as null hypothesis. If both tests point in the same

---

[164] Results are available from the author on request.

[165] See, for instance, Ho (2002) who argued in favor for using ADF and KPSS tests jointly. Thereafter, both results are compared.

direction, the result will be clear. If they contradict each other, the outcomes are said to be uninformative. I stick to this conservative strategy and use a couple of test procedures that are widely applied in the literature. Sarno and Taylor (1998a, b) developed the multivariate augmented Dickey-Fuller panel unit root test, which goes originally back to Abuaf and Jorion (1990). Furthermore, I calculate the pooled ADF test provided by Levin et al. (2002) that can also be used when the cross-sectional dimension exceeds the time dimension. For heterogeneous panels with individual effects, time trends, and common time trends, the test statistic derived by Im et al. (1997) seems to be appropriate. These above mentioned test procedures assume that under the null hypothesis all series have one unit root and are consequently non stationary.

The outcomes for the first differences of share prices and dividends presented in table 5.5 are unambiguous. All tests reject the null hypothesis that all series are I(1), and the Hadri (2000) test cannot reject that all series are stationary. Hence, the results are informative.

Table 5.5: Panel unit root tests for dividends and share prices

I carry out several test procedures with different specifications regarding trends, lags, and the degree of heterogeneity. Note that the Hadri (2000) test allows for unit specific deterministic trends.

Test procedures	First difference of prices		First difference of dividends	
	Test statistic	P-value	Test statistic	P-value
H_0: All series are stationary				
Hadri LM with homogeneous disturbances	-2.939	0.998	-3.285	0.999
Hadri LM with heterogeneous disturbances	-2.993	0.999	-2.783	0.997
H_0: All series are I(1)				
Im test with lag 0 and constant	-5.433	0.000	-7.031	0.000
Im test with lag 1 and constant	-4.605	0.000	-5.039	0.000
Im test with lag 2 and constant	-3.862	0.000	-4.371	0.000
Im test with lag 0, constant, and trend	-5.469	0.000	-6.972	0.000
Im test with lag 1, constant, and trend	-4.727	0.000	-5.014	0.000
Im test with lag 2, constant, and trend	-4.061	0.000	-4.410	0.000
Levin test with lag 0 and constant	-32.343	0.000	-43.493	0.000
Levin test with lag 1 and constant	-27.663	0.000	-30.220	0.000
Levin test with lag 2 and constant	-23.955	0.000	-26.399	0.000
Levin test with lag 0, constant, and trend	-33.055	0.000	-43.796	0.000
Levin test with lag 1, constant, and trend	-28.861	0.000	-30.566	0.000
Levin test with lag 2, constant, and trend	-25.607	0.000	-27.053	0.000
Multivariate Dickey Fuller with lag 1	too large!	0.000	3.69e+06	0.000

The prerequisites of a VAR model are fulfilled. Consequently, I can write the structural panel VAR in the following manner; thereby, to keep the notation simple, I neglect micro- and macro-shocks for the moment.

$$\mathbf{T}\begin{pmatrix}\Delta p_{it}\\ \Delta d_{it}\end{pmatrix} = \mathbf{T}\,\mathbf{\Delta y_{it}} = \mathbf{\Gamma_0} + \sum_{j=1}^{p}\mathbf{\Gamma_j \Delta y_{i(t-j)}} + \mathbf{e_{it}} \tag{5.4}$$

The 2×2 dimensional matrix T captures the impact of current innovations in dividends on stock prices and the other way around. Obviously, this model represents a structural VAR, whose coefficients cannot be estimated directly because not all explanatory variables are predetermined. To allow the determination of the coefficients of the structural form, I permit that current innovations in dividends affect current prices – but not vice versa. This restriction imposed on the matrix T is called Cholesky decomposition and enables to identify the parameters of the structural VAR using the information provided by the reduced form.[166]

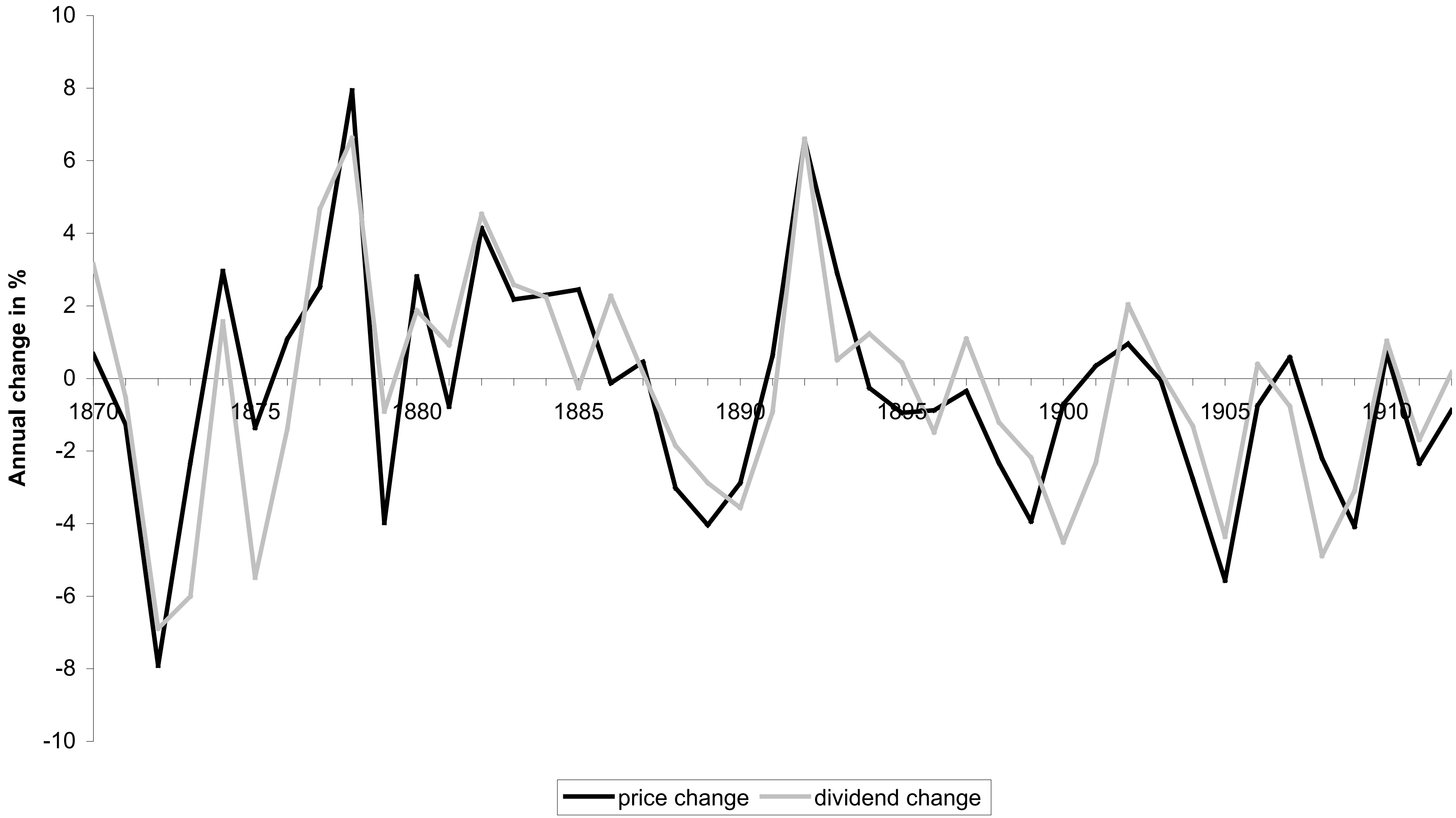
$$\mathbf{T} = \begin{bmatrix} 1 & -\delta \\ 0 & 1 \end{bmatrix} \tag{5.5}$$

The Cholesky decomposition imposes an ordering restriction for the two time series; thereby, I assume that current innovations in fundamentals affect current stock prices – but not vice versa. This restriction is motivated by the theory of informationally efficient markets; hence, one can assume that the stock market anticipates a change in earnings and dividends. In empirical research, this assertion is widespread (see Lee, 1998); thus, the restriction due to the Cholesky decomposition can be justified. In addition, my data structure makes the anticipation of market participants more likely. Note that I collect annual closing prices; hence, these share prices should reflect a large part of the change in dividends paid for the current year. Figure 5.8 highlights that changes in share prices occur prior to subsequent changes in fundamentals.

[166] This proposed recursive system follows Sims (1980) and ensures that the primitive system is exactly identified by imposing the restrictions on the matrix T.

Figure 5.8: The annual change in share prices and dividends from 1870 to 1913

By plotting the annual change in prices and dividends, one can justify the imposed ordering restriction. At a first glance, share prices move prior to dividends.



To derive the standard VAR, the reduced equations, the system (5.4) is pre-multiplied by the inverse of matrix T. The standard form can easily be estimated using OLS or GMM.

$$\mathbf{\Delta y_{it}} = \mathbf{T^{-1}\Gamma_0} + \sum_{j=1}^{p} \mathbf{T^{-1}\Gamma_j \Delta y_{i(t-j)}} + \mathbf{T^{-1}e_{it}} = \mathbf{\Sigma_0} + \sum_{j=1}^{p} \mathbf{\Sigma_j \Delta y_{i(t-j)}} + \boldsymbol{\eta}_{\mathbf{it}} \tag{5.5}$$

The regression of the standard form (5.5) provides coefficients that are related with the parameters of the primitive form as indicated in (5.5). Calculating the variance-covariance matrix of the residuals serves as additional information to identify the coefficients of the structural form. I summarize the interrelation in the equalities (5.6).

$$\begin{aligned} Var\eta_{1it} &= \delta^2\sigma_{e2}^2 + \sigma_{e1}^2 \\ Var\eta_{2it} &= \sigma_{e2}^2 \\ Cov(\eta_{1it}, \eta_{2it}) &= \delta\sigma_{e2}^2 \end{aligned} \tag{5.6}$$

To analyze the dynamics of share prices and dividends triggered by unexpected changes in macroeconomic conditions, I insert macroeconomic shocks in economic growth and inflation into model (5.5). To avoid a debate on exogenity with regard to macro effects, only the unanticipated part of inflation and growth rates are considered in the model. Based on an ARIMA model as discussed in section 5.5.2, individuals expect a specific fluctuation in inflation and economic growth. If these expectations are wrong, the stock market will be hit by an exogenous shock. The 2×1 dimensional vector g_t contains unexpected innovations in inflation and growth rates. Besides analyzing macro shocks, my aim is to evaluate the impact of mergers on the endogenous time series. Hence, by embedding micro-level shocks denoted m_{it}, the effect of mergers can be examined. Note that the dummy variable m_{it} takes the value one if firm i initiate a merger in year t and zero otherwise. The coefficients are stored in the 2×1 dimensional vector M. The panel VAR takes the following shape.

$$\mathbf{\Delta y_{it}} = \mathbf{\Sigma_0} + \sum_{j=1}^{p} \mathbf{\Sigma_j \Delta y_{i(t-j)}} + \mathbf{\Xi g_t} + \mathbf{Mm_{it}} + \boldsymbol{\eta}_{\mathbf{it}} \tag{5.7}$$

As discussed in section 5.4, some mergers can be anticipated by observing inflation rates and the change in nominal capital. Hence, besides regressing model (5.7), I will also consider a specification in which the actual merger m_{it} is replaced by the non-anticipated mergers m_{nt} and the falsely assumed ones m_{nt}.

5.6.3 Determining the lag length p

Determining the lag length p of the VAR should be based on information criterions like Akaike, Hannan Quin, and Schwarz criterion rather than looking at t-tests for significance of the respective lag coefficient. I calculate the information criterions as defined in Hamilton (1994) for my model (5.7); thereby, the lag length is increased from zero to fifteen. Table 5.6 reports the results.

Table 5.6: Information criterions to determine the lag length of the VAR

I calculate the Akaike, Hannan Quin, and Schwarz criterion of the VAR model specified in (5.7). The criterions are calculated with the formulas discussed in Hamilton (1994).

Lags	Information criterions		
	AIC	HQIC	SBIC
0	0.4874	0.4991	0.5182
1	0.3599	0.3795	0.4112
2	0.3324	0.3597	0.4041
3	0.2918	0.3270	0.3841
4	0.2333	0.2763	0.3461
5	0.2124	0.2632	0.3456
6	0.1774	0.2359	0.3311*
7	0.1704	0.2368	0.3446
8	0.1713	0.2455	0.3661
9	0.1643	0.2463	0.3795
10	0.1129	0.2027*	0.3486
11	0.1158	0.2134	0.3720
12	0.1081	0.2135	0.3848
13	0.1107	0.2240	0.4080
14	0.0880	0.2091	0.4057
15	0.0837*	0.2126	0.4220

Consequently, the optimal lag length should be set to six if I follow the Schwarz BIC criterion. Note that the optimal lag length depends heavily on the inclusion of macroeconomic shocks. If unexpected macroeconomic fluctuations were neglected, the optimal lag length would become too large in comparison to the time series dimension. Moreover, microeconomic shocks do not influence the decision regarding the lag structure.

5.6.4 Why is the change in nominal capital not considered in my panel VAR?

Obviously, mergers lead to higher nominal capital if the transaction is financed by issuing new shares, which was very common during the period 1870 to 1913. Therefore, one should at a first glance include the change in nominal capital as third endogenous variable into the panel VAR framework. Lets extent my model by considering the change in nominal capital Δn_{it} and test whether this more complicated model should be used.

$$\begin{bmatrix} \mathbf{\Delta y_{it}} \\ \Delta n_{it} \end{bmatrix} = \mathbf{\Sigma_0} + \sum_{j=1}^{p} \mathbf{\Sigma_j} \begin{bmatrix} \mathbf{\Delta y_{i(t-j)}} \\ \Delta n_{i(t-j)} \end{bmatrix} + \mathbf{\Xi g_t} + \mathbf{M m_{it}} + \boldsymbol{\eta}_{\mathbf{it}} \tag{5.8}$$

Table 5.7 summarizes the outcomes of block F-tests; thereby, I can reject that a change in nominal capital granger causes share prices or dividends. In addition, changes in dividends can clarify the expansion of a company. This finding will be analyzed in a later section that deals with the expansion paths of companies. The results of the reduced form[167] (5.8) show that mergers affect the nominal capital directly (p-value: 0.000). However, the upsurge in nominal capital triggered by a merger does not granger cause share prices or dividends. Hence, for the sake of simplicity, excluding the first difference in nominal capital comes without any loss and reduces the number of parameters to be estimated.

Table 5.7: Granger causality tests

To justify the exclusion of changes in nominal capital from my basic model, I carry out Granger causality tests. Setting appropriate restrictions on model (5.8), F-tests indicate whether the imposed restrictions can be rejected.

Test if respective lagged variables explain endogenous variable		Block F-tests	
Endogenous variable	Predetermined values	F-statistics	p-values
Δp_{it}	$\Delta p_{i(t-j)}\ \forall j = 1,2,\ldots,6$	20.49	0.000
Δp_{it}	$\Delta d_{i(t-j)}\ \forall j = 1,2,\ldots,6$	9.67	0.000
Δp_{it}	$\Delta n_{i(t-j)}\ \forall j = 1,2,\ldots,6$	0.61	0.719
Δd_{it}	$\Delta p_{i(t-j)}\ \forall j = 1,2,\ldots,6$	26.83	0.000
Δd_{it}	$\Delta d_{i(t-j)}\ \forall j = 1,2,\ldots,6$	38.94	0.000
Δd_{it}	$\Delta n_{i(t-j)}\ \forall j = 1,2,\ldots,6$	1.02	0.408
Δn_{it}	$\Delta p_{i(t-j)}\ \forall j = 1,2,\ldots,6$	1.78	0.101
Δn_{it}	$\Delta d_{i(t-j)}\ \forall j = 1,2,\ldots,6$	3.39	0.003
Δn_{it}	$\Delta n_{i(t-j)}\ \forall j = 1,2,\ldots,6$	0.80	0.573
Observations	1262		

However, I put some emphasis on the interrelation between mergers and nominal capital in a later section. But to answer the question whether mergers influence share prices and dividends over a long horizon, focusing on the dynamics of prices and dividends is sufficient.

[167] The results are not reported in a table because they are not of central importance. However, an output table is available on request.

5.7 Empirical findings: Macro and micro-level shocks

5.7.1 Outcomes of the VAR in reduced form – the importance of macro shocks

Table 5.8 summarizes the outcomes of panel VAR models with macroeconomic shocks in reduced form. To compare different estimation procedures, namely OLS and GMM, and the effectiveness of anticipating macroeconomic variables, I run three different models. Using the unexpected inflation and growth rate, OLS and GMM yield similar results. Comparing the model with unexpected variables and the VAR with observed realizations gives the impression that the results are very close to each other. In addition, Granger causality tests underline that prices granger causes dividends and vice versa on the 99% level of confidence. Thus far, only the coefficients of the reduced form are known. To identify the primitive VAR,

Table 5.8: Outcomes of the VAR models in reduced form with macro-shocks

I estimate three different specifications of my VAR model with six lags; thereby, the first two models include unexpected macroeconomic shocks and are estimated by OLS and GMM. The third model incorporates ex post observed realizations of inflation and growth rates.

		OLS estimation unexpected values		GMM estimation unexpected values		GMM estimation ex post observation	
	Variables	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Dependent variable is first difference in share prices	Constant	0.0010	0.859	-0.0062	0.000	-0.0032	0.001
	Δp_{it-1}	0.2077	0.000	0.1100	0.025	0.1208	0.004
	Δp_{it-2}	0.0318	0.181	-0.0226	0.496	-0.0153	0.630
	Δp_{it-3}	0.0085	0.724	-0.0701	0.032	-0.0427	0.163
	Δp_{it-4}	-0.0520	0.026	-0.0882	0.006	-0.0264	0.433
	Δp_{it-5}	0.0107	0.634	-0.0358	0.232	-0.0234	0.371
	Δp_{it-6}	-0.1022	0.000	-0.1319	0.000	-0.1030	0.000
	Δd_{it-1}	0.0365	0.004	0.0364	0.113	0.0355	0.103
	Δd_{it-2}	-0.0515	0.000	-0.0417	0.016	-0.0434	0.011
	Δd_{it-3}	-0.0406	0.001	-0.0303	0.028	-0.0444	0.001
	Δd_{it-4}	-0.0290	0.019	-0.0231	0.155	-0.0398	0.008
	Δd_{it-5}	0.0068	0.578	0.0106	0.410	-0.0025	0.804
	Δd_{it-6}	0.0239	0.035	0.0268	0.139	0.0134	0.437
	Growth rate	0.0078	0.000	0.0142	0.000	0.0036	0.003
	Inflation	-0.0656	0.000	-0.0505	0.000	-0.0631	0.000
	F test	70.42	0.000	197.16	0.000	114.49	0.000
	Adjusted R^2	0.44	-	-	-	-	-
Dependent variable is first difference in dividends	Constant	0.0198	0.112	0.0019	0.339	0.0025	0.201
	Δp_{it-1}	0.5264	0.000	0.4761	0.000	0.4906	0.000
	Δp_{it-2}	0.2248	0.000	0.2017	0.013	0.1801	0.026
	Δp_{it-3}	0.2941	0.000	0.2980	0.000	0.2918	0.000
	Δp_{it-4}	0.2330	0.000	0.2264	0.005	0.2293	0.003
	Δp_{it-5}	0.1679	0.001	0.1900	0.036	0.1953	0.029
	Δp_{it-6}	0.1331	0.006	0.1649	0.004	0.1719	0.003
	Δd_{it-1}	-0.3023	0.000	-0.2975	0.000	-0.2960	0.000
	Δd_{it-2}	-0.2949	0.000	-0.2950	0.000	-0.2906	0.000
	Δd_{it-3}	-0.2754	0.000	-0.2865	0.000	-0.2835	0.000
	Δd_{it-4}	-0.2959	0.000	-0.3158	0.000	-0.3143	0.000
	Δd_{it-5}	-0.1975	0.000	-0.2166	0.000	-0.2175	0.000
	Δd_{it-6}	-0.1074	0.000	-0.1209	0.005	-0.1240	0.004
	Growth rate	0.0128	0.003	0.0121	0.005	0.0109	0.044
	Inflation	-0.0088	0.169	-0.0118	0.075	-0.0141	0.077
	F test	22.40	0.000	42.52	0.000	46.17	0.000
	Adjusted R^2	0.20	-	-	-	-	-
	Observations	1262	-	1227	-	1227	-

the ordering restriction of the Cholesky decomposition summarized in matrix T is needed. By pre-multiplying T and making use of the equalities (5.6), the structural VAR is identified. To illustrate the dynamics captured in my models, a subsequent section focuses on impulse response functions.

5.7.2 Outcomes of the VAR in reduced form – mergers or forecasting errors

Thus far, the estimates of the reduced form point out that macroeconomic shocks are crucial in determining the dynamics of share prices and dividends. To analyze the long-run impact of mergers, three different models are estimated using OLS; thereby, the first model neglects macroeconomic effects – but actual mergers are considered. Table 5.9 contains the outcomes. Without controlling for the economic surroundings in Germany, mergers possess a significant negative effect on real stock prices. In contrast, the second model also incorporates unexpected macroeconomic conditions which terminate the micro-level shock. Because I uncovered – based on the panel probit analysis – that mergers and inflation rates exhibit a strong relation, the last model only includes unexpected mergers. Quantifying surprising mergers is somehow difficult due to the nature of the binary variable. Therefore, based on the assumption that market participants used a model like my probit analysis to anticipate mergers, forecasting errors serve as unexpected shocks. By predicting impending mergers, two mistakes can occur, namely surprising mergers or anticipated but failed mergers. Both forecasting errors may cause different market responses; thus, I separate the two shocks. The error of type one abbreviated m_{pt} denotes the event that a merger is not predicted but executed. Falsely forecasted mergers that do not take place are denoted as error m_{nt} in table 5.9. However, macroeconomic shocks in turn predominate unexpected micro-level effects. Consequently, I draw the conclusion that – apart from the short-term market response provoked by merger announcements – mergers do not influence real share prices and dividend streams in the long-run.

Table 5.9: Outcomes of the panel VAR models in reduced form with mergers

All models are estimated with OLS. GMM leads to quite similar results.

		Model one		Model two		Model three	
	Variables	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Dependent variable is first difference in share prices	Constant	0.0031	0.661	0.0016	0.767	0.0011	0.851
	Δp_{it-1}	0.1807	0.000	0.2069	0.000	0.2072	0.000
	Δp_{it-2}	0.0175	0.560	0.0313	0.188	0.0315	0.185
	Δp_{it-3}	-0.1343	0.000	0.0081	0.737	0.0088	0.714
	Δp_{it-4}	0.0326	0.262	-0.0529	0.023	-0.0529	0.023
	Δp_{it-5}	-0.0557	0.050	0.0101	0.656	0.0108	0.632
	Δp_{it-6}	-0.0962	0.000	-0.1022	0.000	-0.1015	0.000
	Δd_{it-1}	0.0356	0.024	0.0366	0.003	0.0366	0.004
	Δd_{it-2}	-0.0552	0.000	-0.0514	0.000	-0.0516	0.000
	Δd_{it-3}	-0.0170	0.276	-0.0405	0.001	-0.0407	0.001
	Δd_{it-4}	-0.0051	0.744	-0.0286	0.021	-0.0288	0.020
	Δd_{it-5}	0.0295	0.054	0.0070	0.567	0.0068	0.577
	Δd_{it-6}	0.0193	0.176	0.0240	0.034	0.0237	0.037
	Merger m_{it}	-0.0797	0.017	-0.0153	0.563	-	-
	Error I	-	-	-	-	-0.0184	0.557
	Error II	-	-	-	-	0.0153	0.622
	Growth rate	-	-	0.0078	0.000	0.0078	0.000
	Inflation	-	-	-0.0655	0.000	-0.0657	0.000
	F test	11.75	0.000	65.71	0.000	61.59	0.000
	Adjusted R^2	0.11	-	0.44	-	0.44	-
Dependent variable is first difference in dividends	Constant	0.0176	0.169	0.0178	0.163	0.0174	0.176
	Δp_{it-1}	0.5355	0.000	0.5288	0.000	0.5287	0.000
	Δp_{it-2}	0.2134	0.000	0.2262	0.000	0.2257	0.000
	Δp_{it-3}	0.2559	0.000	0.2954	0.000	0.2946	0.000
	Δp_{it-4}	0.2731	0.000	0.2357	0.000	0.2335	0.000
	Δp_{it-5}	0.1618	0.002	0.1699	0.001	0.1699	0.001
	Δp_{it-6}	0.1375	0.005	0.1333	0.006	0.1347	0.006
	Δd_{it-1}	-0.3024	0.000	-0.3026	0.000	-0.3023	0.000
	Δd_{it-2}	-0.2973	0.000	-0.2953	0.000	-0.2950	0.000
	Δd_{it-3}	-0.2715	0.000	-0.2758	0.000	-0.2758	0.000
	Δd_{it-4}	-0.2916	0.000	-0.2971	0.000	-0.2969	0.000
	Δd_{it-5}	-0.1907	0.000	-0.1981	0.000	-0.1980	0.000
	Δd_{it-6}	-0.1061	0.000	-0.1076	0.000	-0.1081	0.000
	Merger m_{it}	0.0292	0.633	0.0448	0.463	-	-
	m_{pt}	-	-	-	-	0.0329	0.649
	m_{nt}	-	-	-	-	0.0434	0.544
	Growth rate	-	-	0.0128	0.003	0.0128	0.003
	Inflation	-	-	-0.0091	0.153	-0.0092	0.152
	F test	22.66	0.000	20.94	0.000	19.61	0.000
	Adjusted R^2	0.19	-	0.20	-	0.20	-
	Observations	1262	-	1262	-	1262	-

To illustrate the complex dynamics captured by my model, the following section discusses the construction of impulse response functions. Thereafter, the bootstrapping approach is highlighted which allows to derive confidence intervals for the impulse response functions.

5.7.3 Impulse response functions

After identifying the structural VAR, I rewrite my model in the vector moving-average representation. A single shock occurs in t = 0. Thus, one can write the dynamic response of both time series in the following manner.

$$\mathbf{\Delta z_{i0}} = \mathbf{T^{-1} e_{i0}} = \mathbf{\Phi_0 e_{i0}} \tag{5.9}$$

$$\mathbf{\Delta z_{i1}} = \mathbf{\Sigma_1 T^{-1} e_{i0}} = \mathbf{\Phi_1 e_{i0}}$$

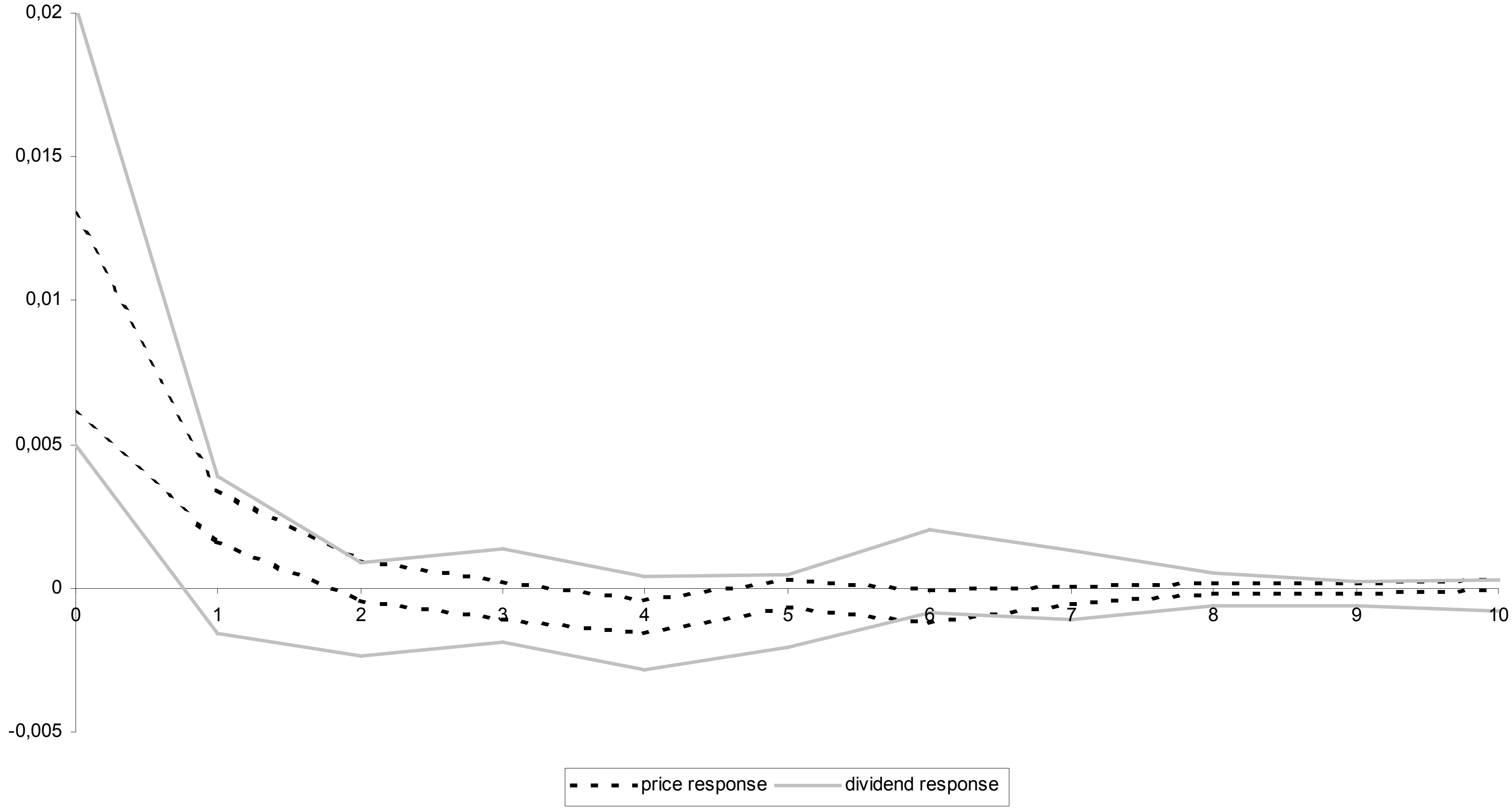
$$\mathbf{\Delta z_{i2}} = \mathbf{\Sigma_1 \Sigma_1 T^{-1} e_{i0}} + \mathbf{\Sigma_2 T^{-1} e_{i0}} = \mathbf{\Phi_2 e_{i0}}$$

To give a simple intuition regarding this calculation, consider that a shock occurs at time t=0. This shock is transferred into the subsequent period by the coefficient matrix Σ_1 that contains the coefficients of the endogenous time series lagged by one period. The next period – two periods after the exogenous shock – is characterized by the shock one period before transferred by the matrix Σ_1. In addition, the initial shock at t=0 is transferred by the matrix Σ_2 that contains the coefficients of the endogenous variables lagged by two periods.

Using this representation, the impulse multipliers included in the matrix Φ_j can be derived for every time horizon. Impulse response functions are the plotted impulse multipliers for the respective point in time j. Figure 5.9 depicts the response of share prices and dividends triggered by an unexpected increase in growth rates of the net national product (NNP) by one percentage point. In line with my expectations, a higher economic activity causes positive responses of share prices and dividends. As described in the next section, I plot the 90% confidence intervals by using the 5% and 95% percentile of the respective bootstrapping

Figure 5.9: Confidence intervals on the 90% level of confidence of impulse response functions for share prices and dividends

Derived from a bootstrapping distribution, I plot the 5% and 95% percentile of share price and dividend responses triggered by an unexpected increase in growth rates of NNP by one percentage point.

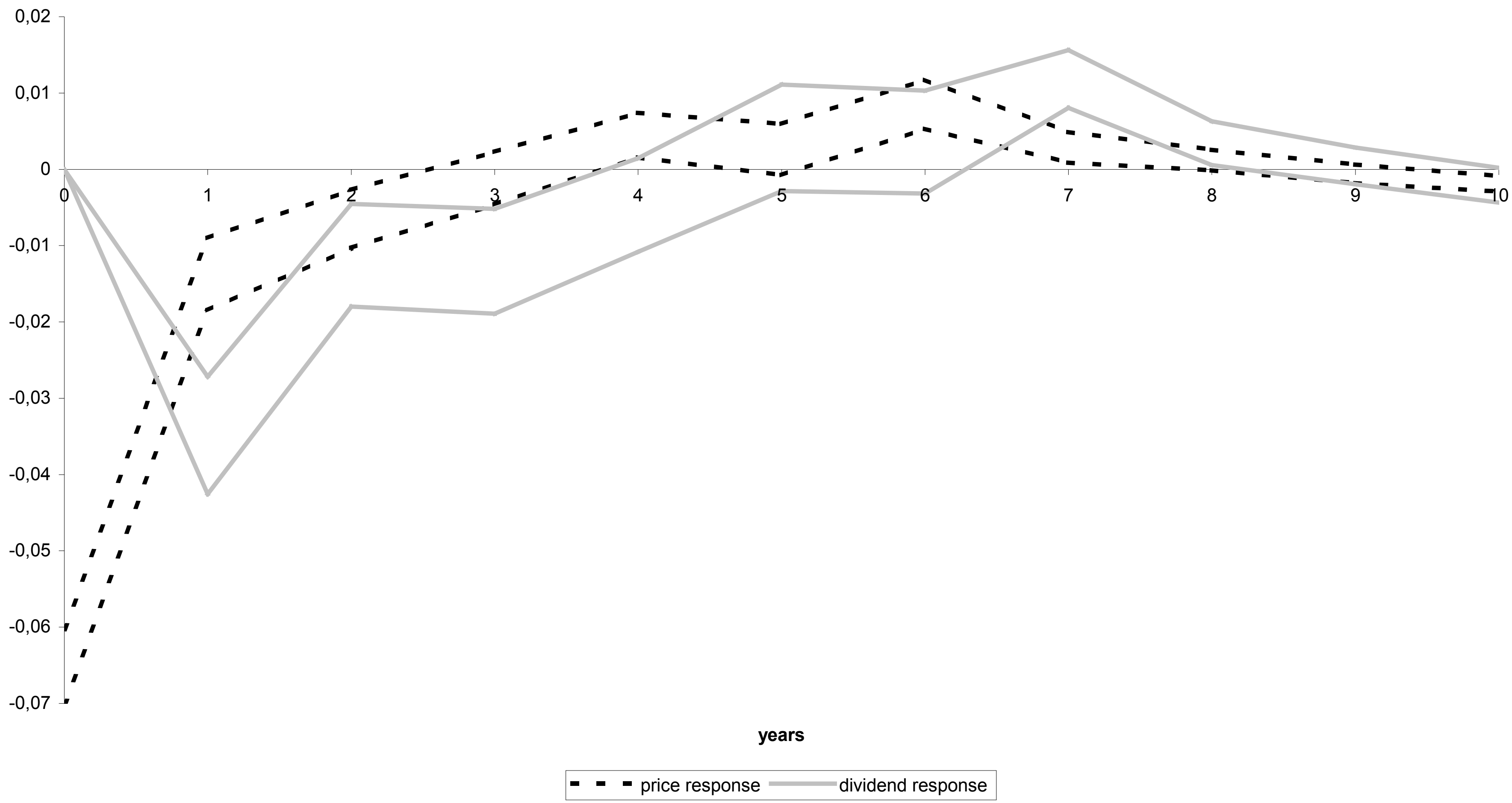


distribution. Caused by the high number of observations, the intervals are relatively small and converge to zero with increasing time horizon. To assess the cumulative effect of a one percentage point increase in growth rates, a bootstrapping distribution is derived, which leads in turn to a 90% confidence interval. Over a ten year period after the positive change in economic growth, share prices increase between 0.60% and 1.39%, whereas dividends upsurge between 0.66% and 1.59%. Generally, the sudden change in macroeconomic conditions is quickly absorb by the market. In addition, share prices and dividends react in a similar manner to changes in economic growth.

In contrast, figure 5.10 indicates an asymmetric reaction when inflation rates fluctuate. A higher inflation by one percentage point has a severe direct impact on share prices. However, inspiring table 5.8 or 5.9 uncovers that inflation does not affect real dividends directly – but dividends are influenced by share price movements. Consequently, share prices fall by 6.55% immediately and the cumulated impact adds up to –6.04% to –7.95% over a ten year period. Also dividends suffer from higher inflation indicated by a negative cumulated effect between –4.97% and –3.30% over a ten year period. However, the change in dividends is less severe than the fall in share prices. Considering the real price-dividend ratio as a measure for valuating a stock, one observes real undervaluation in the presence of unexpected increases in inflation rates, whereas in periods of a reduction in inflation rates, overvaluation results. This finding is in line with Modigliani and Cohn (1979) as well as Ritter and Warr (2002). The real valuation of companies will be precisely discussed in the remainder of this chapter. Moreover, three years after the exogenous shock, the share price response becomes slightly positive until the eighth year – but this counter reaction is not strong enough to outweigh the former decline. This finding points in the direction of Anari and Kolari (2001) that in the long-run a positive impact of inflation on real stock prices should be observed. However, my study stresses the importance of the remarkable negative short-term response of share prices and dividends.

Figure 5.10: Confidence intervals on the 90% level of confidence of impulse response functions for share prices and dividends

Derived from a bootstrapping distribution, I plot the 5% and 95% percentile of share price and dividend responses triggered by an unexpected increase in inflation rates by one percentage point.



5.7.4 Bootstrapping method to derive CI for impulse response functions[168]

To construct confidence intervals for the impulse response functions, one should use bootstrapping methods, rather than searching for an analytical solution. For that purpose, I orient toward Lütkepohl (2000) and use the following bootstrapping approach. After estimating my VAR models, I restore the resulting residuals and use them to reconstruct recursively the share price and dividend time series. Thereafter, the models are re-estimated with the bootstrap time series, and in turn residuals are obtained. I run this process 1000 times and save the impact multipliers Φ_j (j = 1, 2,…,10) on every step. Finally, for every j, one obtains a bootstrap distribution of the respective impact multipliers with 1000 observations. Standard percentile intervals as described by Elfron and Tibshirani (1993) can be used to construct a confidence interval on the 90% level of confidence around the response functions.

5.8 The long-term relation between share prices and dividends

5.8.1 Introduction

Thus far, I know that mergers are more likely in periods that exhibit high inflation rates. In addition, my panel VAR approach uncovered that higher unexpected inflation rates trigger an asymmetric market response. Share prices decline sharply after a sudden increase in inflation – but the reaction of dividends is moderate. If there is a long-term relation between share prices and dividends, the real valuation of companies mitigates during phases of surprisingly high inflation rates. Motivated by these observations, I have the impression that, during the pre-1914 period, mergers are more likely to occur in periods of real undervaluation. Before making a clear statement on the interrelation of mergers and the market valuation, it is essential to discuss methods that identify under- and overvalued companies.

Unfortunately, the often assumed[169] and empirically confirmed finding that the share price dividend ratio[170] is a mean-reverting process cannot be found in my historical data set. Neither unit-root tests applied to individual time series, nor panel based unit-root tests as described in section 5.6.2 can confirm that the price-dividend ratio is stationary. If the price-dividend ratio is not stationary, the ratio does not exhibit a mean-reverting behavior and consequently has no long-term mean. This points out that deviations from an average

[168] The statistical appendix shows some exemplary bootstrapping distributions and the way to obtain them using STATA.

[169] See for instance West (1988) and Kleidon (1986) who propose a simple fad model in which deviations from the long-term average of the price-dividend ratio are seen as perturbations. These perturbations follow a stationary AR(1) process.

[170] In applied research, one often assumes without showing the evidence that the price earnings ratio or the book-to-market ratio have a long-term value and are mean-reverting.

dividend yield are inappropriate to indicate over- or undervaluation of companies. Accordingly, I have to seek more elaborate methods.

In the econometric literature, two highly sophisticated approaches are widely accepted and used. Cointegration analysis enables to detect and estimate the long-run relation between share prices and dividends. A deviation from this equilibrium results in over- or undervaluation of stocks. By inserting an error correction term, which is equal to the deviation from equilibrium, I could estimate a panel VECM and discuss the speed with which share prices and dividends tend back to long-run values. The second method is even more challenging from a technical point of view. By applying advanced time series analysis to share prices and dividends, one can split these time series into a transitory and permanent component. Transitory deviations are seen as a source of short-term under- or overvaluation of companies. The remainder of this chapter discuss both approaches.

5.8.2 Cointegration relation between share prices and dividends

5.8.2.1 Transformation of time series and hidden cointegration

My previous results underline the importance of unanticipated macroeconomic fluctuations on share prices and dividends. Even worse, surprises with regard to inflation rates possess an asymmetric influence. Hence, detecting cointegration becomes a challenging task because one faces the problem of structural breaks due to macro-level shocks. Standard cointegration tests applied to individual time series or the whole panel data set fail in the presence of structural breaks. However, in recent econometric literature cointegration tests that allow for regime shifts were discussed. Consequently, Gregory and Hansen (1996) modified the standard Engle Granger approach to deal with one single structural break. Generally, structural breaks bias the results of usual Augmented Dickey Fuller (ADF) tests as Perron (1989) pointed out. If the points in time at which a shift occurs are known, including a sufficient number of dummy variables into the ADF solves the problem. Unfortunately, I cannot state that structural breaks are known during my investigation period. For unknown structural changes, tests for unit roots exist[171] – but reliable procedures for testing a cointegration relation are still lacking. Especially, if one tries to test for cointegration using the whole panel data set, methods that work in the presence of unknown structural changes have not been developed, as far as I know. I propose a relative simple procedure to correct for structural breaks due to individual or macroeconomic effects. The basic idea is to transform the initial time series so that shocks cannot disturb the long-run equilibrium between modified share prices and dividends.

[171] Perron (1994) used an additive and innovative outlier approach which is to some extent superior in comparison to the procedures proposed by Banerjee et. al (1992) or Andrews and Zivot (1992).

Cointegration between filtered time series can be called hidden cointegration. Transforming time series to get rid of individual or time effects is not really new – but is mainly applied to unbalanced panel data sets. To motivate the transformation of time series, I extend the homogeneous model by allowing company specific effects f_i and time specific shocks d_t.[172] Inserting these effects into my basic model (5.7) leads to the following panel vector autoregression.

$$\mathbf{\Delta y_{it}} = \mathbf{\Sigma_0} + \sum_{j=1}^{p} \mathbf{\Sigma_j \Delta y_{it}} + \mathbf{\Xi g_t} + \mathbf{M m_{it}} + \mathbf{f_i} + \mathbf{d_t} + \mathbf{\eta_{it}} \tag{5.10}$$

Unfortunately, it is not straightforward to quantify individual effects. One may suggest to insert dummy variables and to estimate the respective coefficients. However, using dummy variables to control for company specific effects f_i is misleading caused by the lag structure of the VAR. Note that dummies would be correlated with the lagged dependent variables; hence, it is not possible to observe these effects. To illustrate my argument, consider the case of one specific company, say \`Laurahütte´, which is one of the leading mining companies. If \`Laurahütte´ exhibits negative returns over several prosecuting years in comparison to the sample average, lagged returns are also negative and thus correlated with the dummy variable for the company. This inherent multicollinearity makes a precise assessment of the partial effects impossible. Consequently, I try to consider these effects by transforming my time series, instead of using dummy variables.

By inspiring box plots of first differences in share prices and dividends over time, a pronounced time pattern appears in the case of share prices. However, time effects are less important for the dividend streams. Figure 5.11 illustrates both cases. These time patterns correspond to unexpected macroeconomic shocks that hit the whole economy and can influence – as shown in my panel VAR – stock prices and dividends tremendously. I can eliminate these exogenous macro shocks by mean differencing the initial time series. Love and Zicchino (2003) also favored this procedure for their panel VAR model. To underline that I talk about macroeconomic influences when referring to time shocks, I run a simple regression to explain the time shocks in stock prices and dividends.

[172] Love and Zicchino (2003) used a similar panel VAR approach with individual and time effects.

Figure 5.11: Box plots for first differences in share prices and dividends

In panel A, the first differences in prices exhibit a strong time pattern. This pattern is less important for dividends as depicted in panel B.

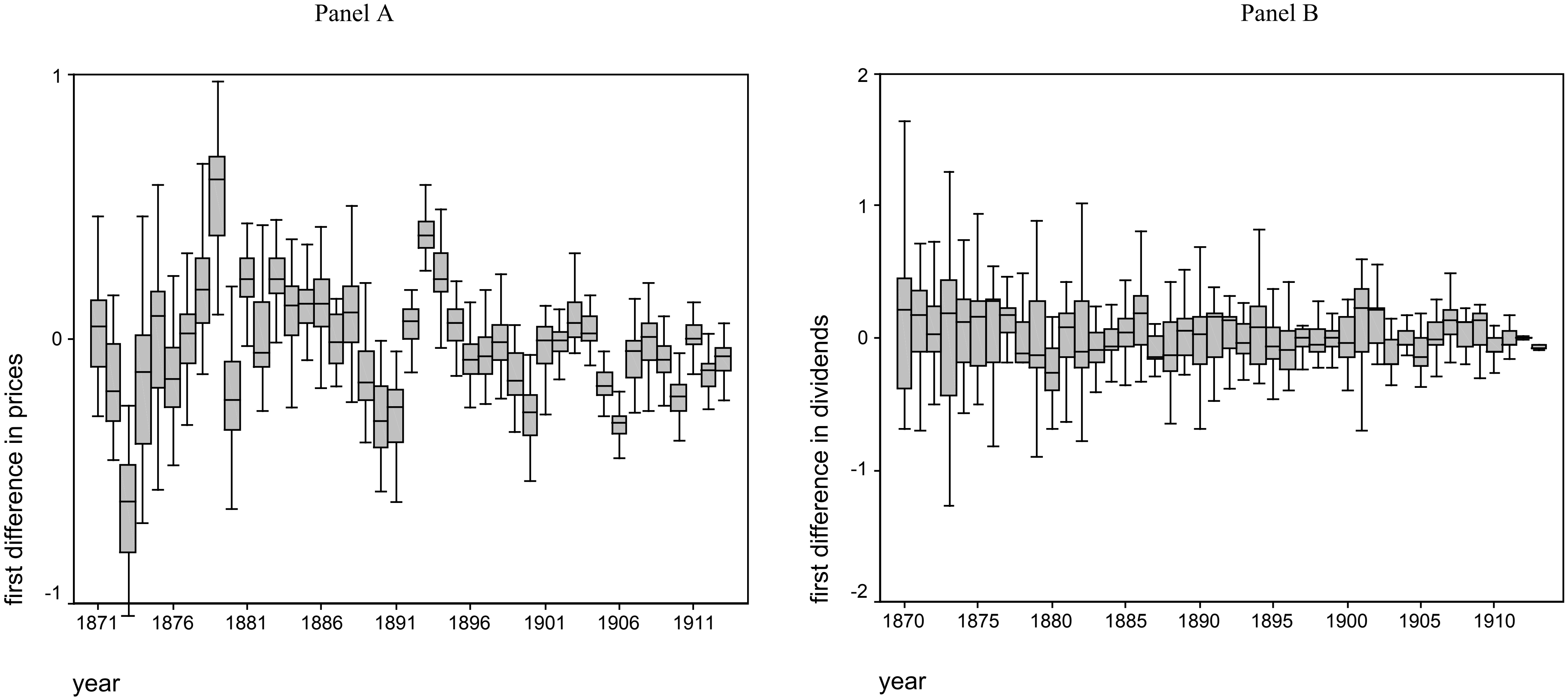


Table 5.10 shows the outcomes together with some regression diagnostics and ADF tests that clearly reject the null hypothesis of I(1) series. If inflation increases, a negative shock on real share prices and dividends results, whereas higher economic growth rates yield positive shocks. Inflation and growth rates are positively correlated with a correlation coefficient of 0.40. These two macroeconomic time series can explain nearly 70% of the variation in share price shocks – but only 20% in the case of dividends. In general, one can conclude that mean differencing the time series eliminates macroeconomic time shocks that hit the whole market.

Table 5.10: Regressions to explain macro-shocks in share prices and dividends

The lag length of the ADF tests is determined by the Schwarz criterion. To test for heteroscedasticity, I use the Breusch-Pagan / Cook-Weisberg procedure. The Ramsey RESET test possesses a high statistical power to detect omitted variable bias and non-linearity.

Dependent variable: macro shocks on prices		
Explanatory variables	Coefficients	P-values
Constant	-0.0263	0.244
Growth rate of NNP	0.0161	0.002
Inflation rate	-0.0739	0.000
Number of Observations	42	
Adjusted R^2	0.69	
F-test	46.34	0.000
Breusch-Pagan / Cook-Weisberg	0.52	0.472
Ramsey RESET	0.23	0.872
Breusch-Godfrey LM (lag 1)	1.04	0.316
Breusch-Godfrey LM (lag 2)	0.75	0.480
Dependent variable: macro shocks on dividends		
Explanatory variables	Coefficients	P-values
Constant	-0.0263	0.244
Growth rate of NNP	0.0207	0.002
Inflation rate	-0.0238	0.022
Number of Observations	42	
Adjusted R^2	0.20	
F-test	6.10	0.005
Breusch-Pagan / Cook-Weisberg	0.59	0.442
Ramsey RESET	1.05	0.384
Breusch-Godfrey LM (lag 1)	0.627	0.434
Breusch-Godfrey LM (lag 2)	1.541	0.228
ADF tests for stationarity		
Variables	Test statistic	Lags (based on Schwarz)
Macro shocks on prices	-3.707***	1
Macro shocks on dividends	-3.945***	1
Growth rate of NNP	-3.778***	1
Inflation rate	-4.313***	1

The second task to eliminate company specific effects denoted f_i is more tricky. Because lagged dependent variables are correlated with these company specific effects, transforming the series is essential – but a simple mean-differencing is not sufficient. Maybe first differencing is already enough to incorporate individual effects. However, a systematically different pattern of dividend growth rates is likely in some newly developing lines of business. For instance, the real estate companies in my sample were founded in the year 1872; hence, these infant companies needed time to generate enough revenue and profits for issuing dividends. Therefore, I prefer an appropriate transformation to control for entity specific circumstances. For that purpose, I apply the Helmert's transformation as proposed by Arrelano and Bover (1995) to my data set. This transformation is also common in applied research on dynamic panel data, see, for instance Bond and Meghir (1994). Moreover, Love and Zicchino (2003) applied this procedure to analyze a panel VAR with individual effects. The Helmert's procedure transforms the time series in levels by subtracting the future expected value from the current value of the variable.

$$\mathbf{y}_{\mathbf{it}}^{*} = \left(\frac{T-t}{T-t+1} \right)^{0.5} \left(\mathbf{y}_{\mathbf{it}} - \frac{1}{T-t} \sum_{j=1}^{T} \mathbf{y}_{\mathbf{i(t+j)}} \right) \tag{5.11}$$

Obviously, using these transformed variables in my panel VAR model violates the assumption of weak exogenity because the variables incorporate future information. Thus, transformed variables are not predetermined. To estimate the VAR with modified series, one has to apply the GMM procedure as thoroughly discussed in Arellano and Bover (1995). Thereby, the non transformed lagged variables serve as instruments for the modified variables.

The following step is to apply individual or panel cointegration tests to the untransformed time series, the mean-differenced series without macroeconomic effects, and the sequences after the Helmert's method.

5.8.2.2 Individual tests for cointegration

Without concerning the macroeconomic time shocks, finding cointegration between share prices and dividends is much easier because structural breaks disappear after mean differencing. Table 5.11 reports the trace statistics of Johansen tests for cointegration; thereby, an intercept and time trend is included. As thoroughly discussed above, I use the VAR with lag length six and try to uncover the long-run relation. The Johansen procedure is applied to the original time series containing time shocks and the time series after eliminating individual and time effects. While I can reject ten times that there is no cointegration considering the unmodified time series, I find cointegration in 17 cases after eliminating individual and time effects. Thus, uncovering the cointegration relation between share prices and dividends is more convincing based on transformed time series.

However, the results of the Johansen procedure should be interpreted carefully because I have to deal with relatively short time periods. Of course, it is possible to improve the finite sample properties by correcting the trace or the critical value. Cheung and Lai (1993) as well as Reimers (1991) provided useful corrections for small samples. If I apply these corrections to my results in table 5.11, detecting cointegration will become impossible. Note that the correction factor used by Cheung and Lai (1993) to increase the critical boundary of the trace statistic is n/(n-kp). Thereby, n denotes the sample size of individual time series, which is below 43, k represents the number of estimated coefficients, and p is the lag length. Because the optimal lag length and hence the number of coefficients is determined by my panel based VAR, the correction factor becomes even negative. Therefore, one should define an optimal VAR specification on the level of individual time series. Using the Schwarz SBIC information criterion, I can specify a VAR with only one lag. Obviously, this reduces the correction factor for the critical value tremendously. Hence, the critical trace is only multiplied with about 1.13. However, when one now assesses the possibility to reject the null hypothesis using the corrected critical values, the picture is quite similar. In the case of unmodified time series, the null hypothesis can be rejected nine times, whereas after Helmert's transformation and mean differencing 33 time series exhibit cointegration. Generally, eliminating time effects and with that macroeconomic shocks facilitates to detect cointegration. Because the Johansson procedure has a sequential character, it is usually recommended[173] to test the null hypothesis with increasing cointegration rank untill the null hypothesis cannot be rejected anymore. Table 5.11 provided the outcomes for the null hypothesis that cointegration does not exist; hence, the rank is equal to zero. However, testing for one cointegration relation uncovers that the null hypothesis for rank equal to one cannot be rejected regardless which type of time series is considered.[174]

Consequently, using transformed instead of original time series allows to justify a cointegration analysis. Caused by the fact that I work with a panel VAR approach, testing for cointegration for every cross-sectional unit is just a first step. Thus, the following section discusses several strategies to identify cointegration in panels. Unfortunately, this area of research is still `under construction' and represents an infant branch of applied econometrics. Taking this fact into account, I prefer to use a couple of different methods to mitigate the scale of uncertainty inherent with every test procedure.

---

[173] See, for instance Harris and Sollis (2003).

[174] To save space, I skip the output tables.

Table 5.11: Johansen tests applied to individual time series

The null hypothesis is no cointegration against the alternative that share prices and dividends have a long-term equilibrium. I only report the trace statistic, which is common in applied work. One star indicates that the null hypothesis can be rejected. The null hypothesis states that cointegration between share prices and dividends does not exist.

Company code	With individual and time effect		Transformed time series	
	Panel VAR	Individual VAR	Panel VAR	Individual VAR
1	15.89	8.74	31.36*	29.39*
2	27.56*	8.94	29.73*	18.82
3	8.83	12.81	19.60*	33.43*
4	55.57*	16.95	45.97*	57.24*
5	8.52	15.90	17.84	31.38*
6	24.36*	40.78*	24.86*	70.55*
7	7.19	8.21	17.24	35.73*
8	21.30*	29.11*	26.44*	31.64*
9	17.97	2.68	18.11	23.01*
10	23.11*	28.49*	15.10	45.29*
11	15.80	16.65	24.47*	22.69*
12	12.96	15.25	31.38*	40.66*
13	8.71	24.28*	43.33*	40.25*
14	17.38	31.55*	22.02*	43.69*
15	21.90*	21.17	30.18*	39.08*
16	8.08	9.05	15.08	31.10*
17	7.95	22.22	18.59*	52.17*
18	7.88	18.57	31.16*	24.04*
19	16.74	9.67	8.40	44.65*
20	9.85	15.81	34.22*	24.49*
21	17.30	29.41*	26.80*	48.85*
22	16.47	17.33	21.15*	36.89*
23	11.15	6.01	20.75*	22.66*
24	13.28	7.74	43.87*	12.55
25	21.44*	17.52	29.60*	32.23*
26	17.72	9.49	52.05*	47.75*
27	14.24	6.82	27.71*	36.37*
28	9.50	8.93	18.40*	23.35*
29	9.37	30.02*	12.45	52.31*
30	8.66	11.81	16.56	30.39*
31	18.60*	7.52	31.89*	31.82*
32	20.27*	32.92*	29.17*	54.14*
33	14.21	6.12	35.12*	54.53*
34	22.06*	23.81*	36.43*	30.02*
35	9.82	17.35	19.36*	35.49*
	Critical values with intercept and trend in VAR			
	Six lags	One lag	Six lags	One lag
Uncorrected	18.17	19.96	18.17	19.96
Cheung / Lai	-	22.59	-	22.59

5.8.2.3 Testing for cointegration in panels

As mentioned in the discussion on panel unit root tests, the cointegration tests can be improved when applied to the whole data set. Maybe the simplest test procedure was developed by Larsson, Lyhagen, and Lothgren (2001) and is based on my former results of the Johansen test statistics for individual series. They proposed to calculate the average of the individual trace statistics. Thereafter, a standardized likelihood ratio statistic is derived using the moments of the asymptotic trace statistic. Fortunately, these moments are presented in Larsson et al. (2001);[175] hence, I get the following test statistic.

$$\frac{\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} tr_i - E(Z_{2-r})\right)}{\sqrt{Var(Z_{2-r})}} \sim N(0,1) \tag{5.12}$$

Because only two variables are included in the VAR, I use the moments for the asymptotic trace statistic Z_{2-r}; thereby, r denotes the cointegration rank used in the null hypothesis. N (=35) represents the number of cross-sectional units and tr_i are the individual trace statistics. To derive the test statistic, I use the individual trace statistics as shown in table 5.11 for the VAR with six lags. This seems to be appropriate because I try to detect cointegration on the panel level. The test statistic developed by Larsson et al. (2001) is a one tailed test; thus, if the test statistic is larger than the respective standard normal quartile, one can reject the null hypothesis. The test statistic reaches 18.16 for unmodified time series and 37.15 for transformed time series; the null hypothesis of no cointegration is obviously rejected in both cases. Because the Johansen procedure should be executed sequentially, I now change the null hypotheses. Thus, I try to reject that there is cointegration of rank one in the data. Unfortunately, the null hypothesis is rejected for modified time series with a test statistic of about 23.50. Nevertheless, with a test statistic of –0.68 obtained from unmodified series, one cannot reject the null hypothesis. Thus, the result is ambiguous with regard to the pros and cons of transformations.

To obtain a clearer picture, I utilize the methodology developed by Pedroni (1999) who tests for cointegration in heterogeneous panels.[176] For computational simplicity, I reduce the degree of heterogeneity a bit in the sense that I stick to the same lag structure when shifting to the next cross-sectional unit. This does not affect my results because the panel is more homogeneous than assumed by Pedroni's (1999) approach. I carry out different specifications of the following regression; thereby, D_i denotes a dummy variable for the respective cross-sectional unit.

---

[175] See table 1 on page 114.

[176] A STATA 8.0 program is available from the author on request.

$$p_{it} = \sum_{i=1}^{N} \alpha_i D_i + \sum_{i=1}^{N} \delta_i D_i t + \sum_{i=1}^{N} \beta_i D_i d_{it} + \upsilon_{it} \qquad t = 1,...,T, i = 1,...,N \tag{5.13}$$

Thus, I allow that the intercept and the slope coefficient may vary across companies, which implies a variety of long-run equilibriums between share prices and dividends. In a second test, I restrict the possibility to have different intercepts. Furthermore, I regress with and without company specific individual time trend, which is represented by the second term in equation (5.13). The residuals of the regression of dividends on share prices are used to derive the test statistics listed in Pedroni's (1999) first table. Thereafter, I run a second regression in first differences and calculate the long-run variance of the residuals and the simple variance.

$$\Delta p_{it} = \sum_{i=1}^{N} \gamma_i D_i \Delta d_{it} + u_{it} \qquad t = 1,...,T, i = 1,...,N \tag{5.14}$$

Both statistics, namely the long-run and the simple variance of the residual u_{it} are nuisance parameters for the test statistics. The Newey-West (1987) estimator with a bandwidth of five determines the long-run variance. For the non-parametric test statistics, I estimate the first order autoregression of the residuals υ_{it} (5.15).

$$\upsilon_{it} = \sum_{i=1}^{N} \delta_i D_i \upsilon_{it-1} + \omega_{it} \qquad t = 1,...,T, i = 1,...,N \tag{5.15}$$

In contrast, to derive the parametric test statistics, one has to carry out the following estimation procedure.

$$\upsilon_{it} = \sum_{i=1}^{N} \vartheta_i D_i \upsilon_{it-1} + \sum_{i=1}^{N} \kappa_i D_i \sum_{k=1}^{K} \Delta \upsilon_{it-k} u_{it} + \nu_{it} \quad t = 1,...,T, i = 1,...,N \tag{5.16}$$

I set K equal to four for all cross-sectional units. This makes the computation easier and does not affect the outcomes. The variances of the residuals from the two autoregressions are also used as nuisance parameters to derive the test statistics.

After identifying each nuisance parameter, I construct the seven test statistics listed in table 1 (see Pedroni, 1999). Based on Pedroni's (1999) adjustment terms (see table 2), I standardize the test statistics which are then standard normally distributed. Table 5.12 summarizes the outcomes for the initial time series and the modified ones. To reject the null hypothesis of no cointegration the panel ν statistic have to be larger than the critical value of the right tail of the standard normal distribution. In all other cases, the left tail of the standard normal distribution is used to reject the null hypothesis. Allowing an individual intercept and a company specific deterministic trend in the regression (5.13), all seven test statistics reject the null if time shocks are eliminated. Note that the presence of macroeconomic shocks yields structural breaks of the long-run relation between share prices and dividends. Consequently, neglecting time effects facilities to detect cointegration. So the transformation of the time series by mean-differencing can be justified; however, the Helmert's transformations seems to be less effective. As mentioned above, focusing on returns and correspondingly first differences may be sufficient to eliminate company specific effects. Justified by the tests for cointegration, I now turn to the estimation of the long run equilibrium.

Table 5.12: Results of the Pedroni (1999) procedure

I summarize the outcomes of the Pedroni (1999) test statistics and their p-values. Note that for the first test statistic the right tail of the standard normal distribution is relevant, whereas for the others the left tail have to be applied. P-values are set in parentheses.

	Test statistics for initial time series					
	Homogeneous intercept		Individual intercept		Ind. Intercept and trend	
Panel ν statistic	-4.58	(1.000)	-6.65	(1.000)	1.64	(0.051)
Panel ρ statistic	-1.16	(0.123)	1.29	(1.000)	0.69	(0.755)
Panel t statistic (non-parametric)	-46.53	(0.000)	-55.05	(0.000)	-88.63	(0.000)
Panel t statistic (parametric)	-1419.52	(0.000)	-231.92	(0.000)	-429.49	(0.000)
Group ρ statistic	1.36	(0.913)	3.74	(1.000)	2.74	(0.997)
Group t statistic (non-parametric)	-72.45	(0.000)	-70.78	(0.000)	-95.66	(0.000)
Group t statistic (parametric)	-78.66	(0.000)	-77.14	(0.000)	-110.63	(0.000)
	Test statistics for time series without macroeconomic time shocks					
	Homogeneous intercept		Individual intercept		Ind. Intercept and trend	
Panel ν statistic	-4.58	(1.000)	-6.66	(1.000)	46.39	(0.000)
Panel ρ statistic	-0.56	(0.288)	1.96	(0.975)	-5.61	(0.000)
Panel t statistic (non-parametric)	-43.90	(0.000)	-51.71	(0.000)	-135.71	(0.000)
Panel t statistic (parametric)	-1347.69	(0.000)	-397.44	(0.000)	-626.07	(0.000)
Group ρ statistic	1.36	(0.913)	3.69	(1.000)	-2.89	(0.002)
Group t statistic (non-parametric)	-69.57	(0.000)	-68.10	(0.000)	-158.13	(0.000)
Group t statistic (parametric)	-76.94	(0.000)	-74.85	(0.000)	-182.70	(0.000)
	Test statistics for time series after Helmert's transformation					
	Homogeneous intercept		Individual intercept		Ind. Intercept and trend	
Panel ν statistic	-4.58	(1.000)	-6.58	(1.000)	12.42	(0.000)
Panel ρ statistic	-3.77	(0.000)	1.12	(0.869)	-5.17	(0.000)
Panel t statistic (non-parametric)	-59.67	(0.000)	-48.33	(0.000)	-128.00	(0.000)
Panel t statistic (parametric)	-3865.41	(0.000)	-391.31	(0.000)	-975.14	(0.000)
Group ρ statistic	-1.71	(0.044)	3.24	(0.999)	-3.27	(0.000)
Group t statistic (non-parametric)	-89.72	(0.000)	-60.78	(0.000)	-162.51	(0.000)
Group t statistic (parametric)	-93.23	(0.000)	-70.16	(0.000)	-182.27	(0.000)

5.8.2.4 Estimating the cointegration vector between prices and dividends

The series after eliminating time shocks strongly exhibit co-movement confirmed by the individual tests as well as the panel cointegration tests. Because these modified time series represent only components of the original time series, one can call the long-run relation hidden cointegration. An additional advantage of using modified series is that the issue how to correct for heterogeneity over time is avoided. This facilities to estimate the long-run equilibrium. Besides this merely technical advantages, my understanding of a long-run equilibrium in an economy corresponds to a state not affected by severe exogenous disturbances. I regard and model unexpected macroeconomic shocks as sources for short-term responses of share prices and dividends. Hence, one should not estimate the long-term interrelation between stock prices and dividends based on the time series `polluted' by macroeconomic shocks. Furthermore, it is worthwhile mentioning that one can hardly observe a difference between time series transformed by mean-differencing and those modified by the Helmert's procedure. This stresses the predominance of macroeconomic shocks compared to mergers as micro level shocks.

The current literature like Pedroni (2000) favored the usage of between group estimators in comparison to within dimension estimators. Applying the group means estimator, which belongs to the between group estimators, as proposed by Pedroni (2000) delivers panel cointegration and individual cointegration vectors. In contrast, within dimension estimators force all companies to have the same cointegration vector under the alternative hypothesis. These estimators only reveal the sample mean of the underlying long-term relation. Hence, I use a group means estimator and utilize a dynamic OLS procedure (see Pedroni, 2001). Accordingly, I want to obtain estimates for the following panel regression model. The between group estimator permits individually different long-term relations.

$$p_{it} = \alpha_i + \beta_i d_{it} + u_{it} \quad t = 1,...,T, i = 1,...,N \tag{5.17}$$

The strict exogenity of the explanatory variable is not fulfilled; hence, OLS is asymptotically biased. Inserting a sufficient number of lead and lags of first differenced regressors eliminates the endogenity bias.

$$p_{it} = \alpha_i + \beta_i d_{it} + \sum_{k=-K_i}^{K_i} \gamma_{ik} \Delta d_{i(t-k)} + u_{it} \quad t = 1,...,T, i = 1,...,N \tag{5.18}$$

Following Pedroni's (2001) discussion, the standard dynamic ordinary least squares (DOLS) estimator can be used to obtain estimates for β_i. Inference should be based on the standard heteroscedasticity and autocorrelation consistent (HAC) estimation of the variance covariance matrix of the augmented regression (5.18). The disadvantage of this technique is the loss of many observations caused by the inclusion of leads and lags. Table 5.13 shows the results of β_i with varying length K_i of the two sided filter. Unfortunately, the individual DOLS estimates are based only on a few observations - about 36 observations over time. Consequently, I should estimate a group mean vector as thoroughly discussed by Pedroni (2001). The DOLS group mean estimator is simply the average individual conventional DOLS estimate. It is also straightforward to derive a confidence interval for the group mean estimator having obtained the individual HAC corrected t statistics. I define the DOLS group mean point estimate and the corresponding t-values as follows. Similar formulas are applied to the estimation of the intercept.

$$\hat{\beta}_{GD} = N^{-1}\sum_{i=1}^{N}\hat{\beta}_{DOLS,i} \tag{5.19}$$

$$t_{\hat{\beta}_{GD}} = N^{-1/2}\sum_{i=1}^{N}t_{\hat{\beta}_{DOLS,i}} \tag{5.20}$$

The resulting group panel estimator for the slope coefficient reaches 0.5266 (t-value: 11.72), whereas the intercept is –0.0208 (t-value: -7.14). By inspiring a simple scatter plot with pooled data and the fitted curve underlines that there is little heterogeneity in the individual long-term relation between share prices and dividends. Figure 5.12 provides the evidence.

Table 5.13: Estimated coefficients of the long-run relation using DOLS

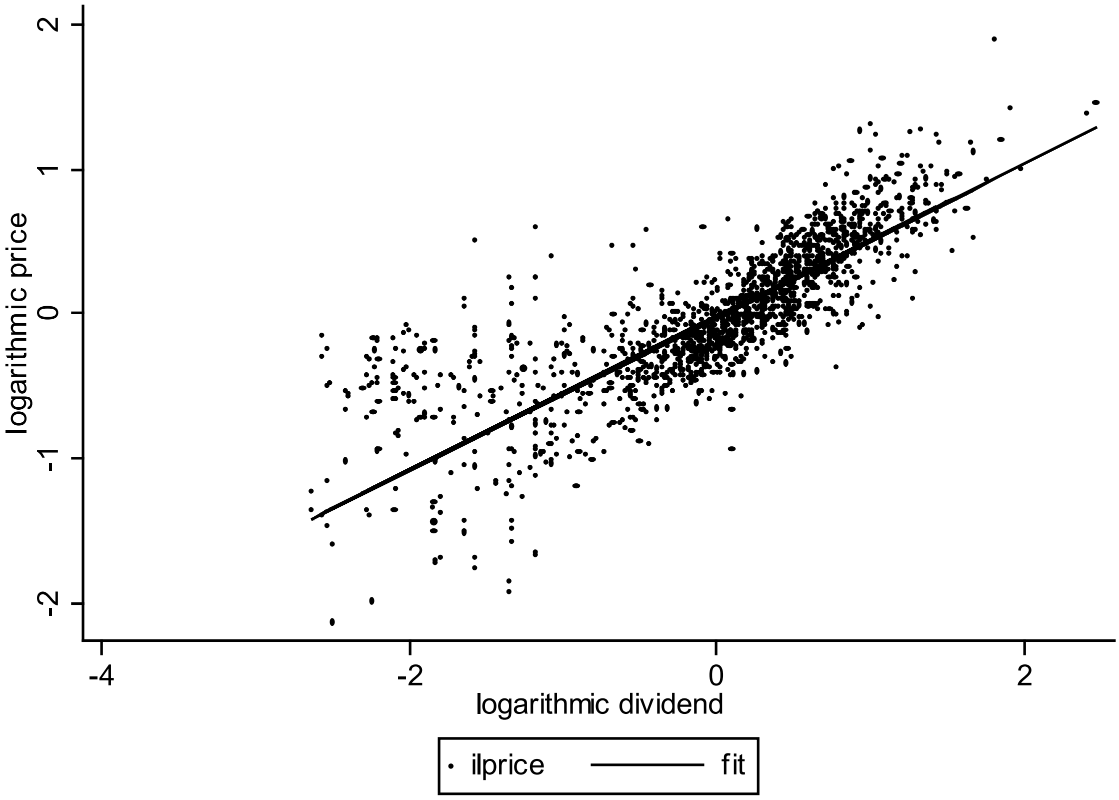
I estimate by applying DOLS the individual coefficients of long-term equilibriums between share prices and dividends with varying numbers of leads and lags.

Company code	$K_i = 1$		$K_i = 2$		$K_i = 3$	
	α_i	β_i	α_i	β_i	α_i	β_i
1	0.0485	0.1104*	0.0485	0.1271**	0.0491	0.1108
2	-0.0769**	0.7569***	-0.0834***	0.7993***	-0.0847***	0.8080***
3	0.0248	0.7728***	-0.0025	0.8253***	-0.0168	0.8525***
4	-0.0008	0.5748***	-0.0234	0.6126***	-0.0244	0.6120***
5	-0.1283***	0.2975	-0.1560***	0.1813	-0.1499***	0.2683
6	-0.0475**	1.2654***	-0.0656***	1.4091***	-0.0630***	1.4327***
7	-0.0738***	0.6990***	-0.0732***	0.6919***	-0.0771***	0.6460***
8	-0.1013***	0.5713***	-0.1461***	0.9463***	-0.1571***	1.0848***
9	-0.0431	0.7387***	-0.0579	0.7788***	-0.0594	0.7898***
10	-0.3694***	-0.3533	-0.4425***	-0.5858*	-0.5144***	-0.7820**
11	-0.1196***	0.7376***	-0.1298***	0.8159***	-0.1288***	0.8816***
12	-0.0328	0.5598***	-0.0400	0.5630***	-0.0710	0.6869***
13	0.1031**	0.3861***	0.1087***	0.4867***	0.1065***	0.5341***
14	-0.0821**	0.4621***	-0.0852**	0.5059***	-0.0910*	0.5310***
15	-0.1528***	0.6100***	-0.1558***	0.6064***	-0.1372***	0.5809***
16	0.2907***	0.3757***	0.2990***	0.3698***	0.2863***	0.3484***
17	0.0704	0.8247***	0.0535	0.8532***	0.0256	0.8850***
18	-0.0646	0.2630***	-0.0380	0.2967***	-0.0137	0.3309***
19	0.2088***	0.2883***	0.2413***	0.2188***	0.2858***	0.1450*
20	-0.2106*	0.5880***	-0.0203	0.7631***	0.1904**	0.9639***
21	-0.1502***	0.9115***	-0.1411***	0.9508***	-0.1213**	1.0119***
22	-0.2375	0.1065	-0.2994*	0.0671	-0.3933**	0.0074
23	0.1928***	0.4095***	0.1656	0.4656***	0.2219*	0.4296**
24	-0.2111***	0.3542***	-0.1734**	0.4078***	-0.1112	0.5338***
25	-0.0426	0.2754***	-0.0177	0.3492***	0.0040	0.4178***
26	-0.1469***	0.3734***	-0.1422***	0.3491***	-0.1420***	0.2770***
27	0.0578*	0.5973***	0.0327	0.6423***	0.0376	0.6472***
28	-0.0850**	0.2230***	-0.0796*	0.2417***	-0.0769*	0.2521***
29	0.2967***	0.4780***	0.1268	0.6685***	0.0970	0.7144***
30	-0.3625***	0.3472***	-0.4058***	0.3088***	-0.4222***	0.3013***
31	-0.1815***	0.5062***	-0.2002***	0.6174***	-0.1993***	0.6565***
32	-0.0464	0.7734***	-0.0596	0.7842***	-0.0412	0.7706***
33	-0.0513	0.6356***	-0.0350	0.6251***	-0.0054	0.5883***
34	0.2060	0.3925	0.3686	0.2266	-0.1494	0.0563
35	-0.1228**	0.5594***	-0.1487***	0.5726***	-0.1664***	0.5583***

(* indicates significance on the 10%, ** on the 5%, and *** on the 1% level of significance)

Figure 5.12: Scatter plot with pooled observations of share prices and dividends

I plot share prices against dividends and include Pedroni's (2001) DOLS between-dimension estimator for the intercept and the slope coefficient.



In contrast to other empirical studies based on `modern´ data, the slope coefficient is relatively low. Nevertheless, a positive relation between share prices and dividends can be confirmed; thereby, a cointegration vector of (1, -1) as implicitly suggested by the simple fad model discussed in West (1988) is not confirmed. Note that West (1988) assumed that the logarithmically transformed price dividend ratio is a mean reverting process. Hence, the ratio must be stationary which implies an cointegration relation between share prices and dividends with the specific (1, -1) cointegration vector.

The knowledge with regard to the long-term equilibrium can be exploited in my VAR models by inserting error correction terms and revealing the speed of adjustment.[177] However, my major aim is to evaluate whether share prices are over- or undervalued in real terms.

5.8.2.5 Interrelation between merger activity and real valuation of stocks

To illustrate whether stock prices deviate from their long-term equilibrium, figure 5.13 depicts the actual changes of share prices and the fundamentally justified fluctuations. At a

[177] The VAR should then be based on modified time series.

first glance, the period around 1900 exhibited a real undervaluation. A merger wave that stared in the year 1898 can be confirmed by my data. Furthermore, for the US case, Banerjee and Eckhard (2001) found similar results. To illustrate the interrelation between the merger wave and the real valuation of companies, figure 5.14 depicts the aggregated deviation from fundamentally justified levels and the merger activity during the period 1898 to 1911. I quantify the merger activity as the number of companies executing at least one merger in the respective year relative to the total number of companies in my sample. This relative merger activity is expressed in per cent. By adding up the deviations of actual share prices from fundamentally justified levels, I get a cumulated deviation. Plotting this cumulated figure against the merger activity is motivated by the fact that the decision to merge is a long-term decision. By looking at figure 5.13, short periods of undervaluation were common – but did not lead to mergers. In contrast, the period from 1898 to 1911 can be characterized as a longer phase of real undervaluation that drives share prices further away from fundamentals over several years.

Figure 5.13: Fundamentally justified and actual share price changes based on an equally weighted portfolio

Plotting the justified and the actual change in share prices highlights periods of real over- or undervaluation.

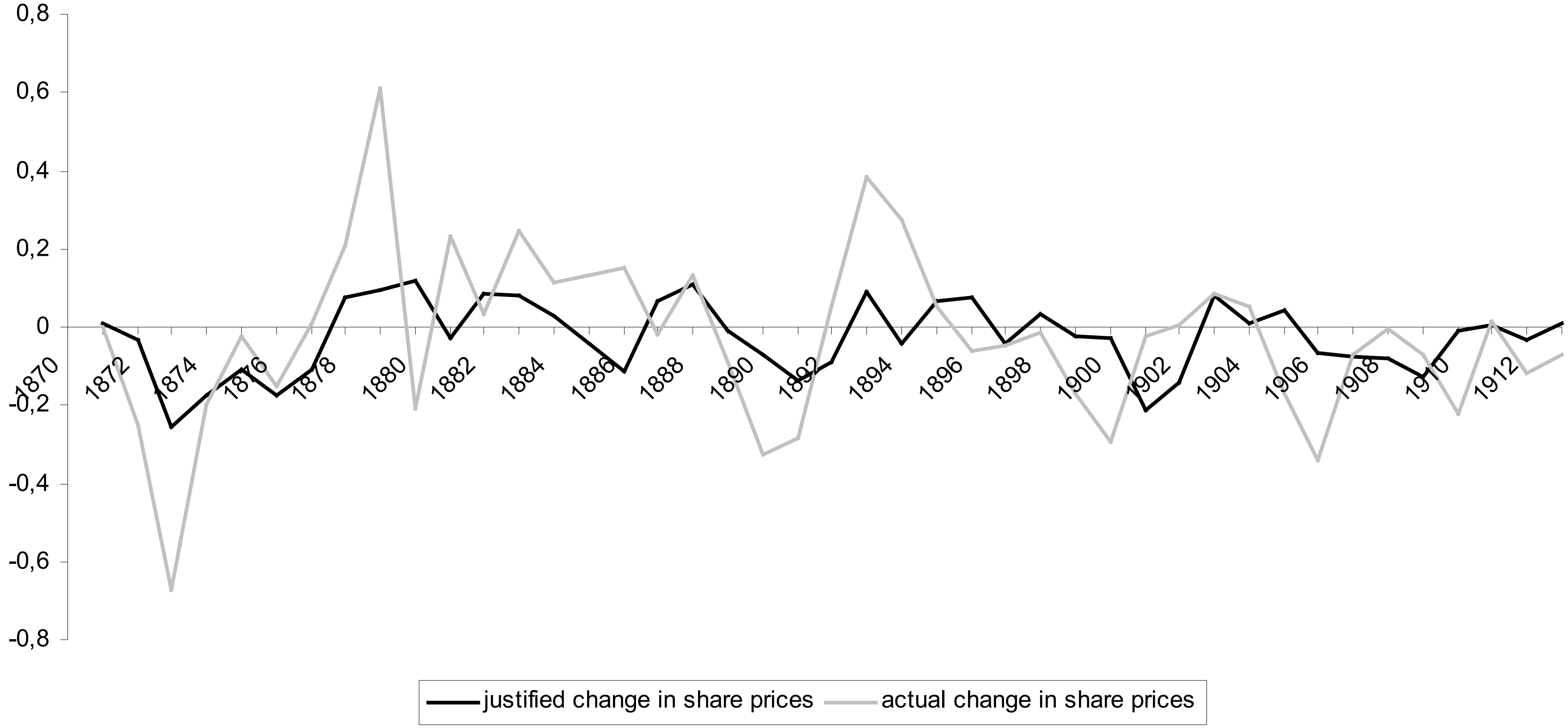


Figure 5.14: Interrelation between the merger activity and the aggregated real undervaluation of equity

By inspiring the merger activity and the cumulated undervaluation of stocks, a positive interrelation becomes apparent.

5.8.3 Decomposing the transitory and permanent components of time series

5.8.3.1 The econometric procedure

An alternative approach to assess whether share prices deviate from their fundamentals respectively from a long-run level is to decompose the time series into permanent and transitory components. To measure whether stocks are over or undervalued, I concentrate on the transitory components. This is based on the imagination suggested by the simple fad model (see West, 1988) that the long-term equilibrium is influenced by short-term perturbation which are stationary by nature. Because I analyze panel data, the decomposition is applied to i individual time series. Each time series possesses a transitory component and a permanent component and, hence, one can express the vector of the three time series z_{it} in the following manner.

$$\begin{pmatrix} p_{it} \\ d_{it} \\ n_{it} \end{pmatrix} = \mathbf{z_{it}} = \mathbf{x_{1,it}} + \mathbf{x_{2,it}} \tag{5.21}$$

In contrast to my analysis above, z_{it} also includes the logarithmically transformed nominal capital n_{it}. Note that including three instead of two time series enables to split the time series into three components.[178] Furthermore, the time series are not transformed. The 3×1 vector denoted $x_{1,it}$ stands for the long-term component; thus, it is required that this component is integrated of order one I(1). Whereas $x_{2,it}$ describes the transitory component and is therefore stationary. Note that this features enable to separate the time series later.

What makes this technique so fascinating besides its pure econometric \`beauty´? First, whether share prices, dividends, and nominal capital are cointegrated is not a prerequisite of the decomposition.[179] Second, no observations are lost due to estimation procedures compared to the dynamic ordinary least squares (DOLS) approach. Finally, structural breaks cannot influence the outcomes – but may affect my cointegration results. Therefore, it seems to be worthwhile to thoroughly discuss this elaborate econometric device.

In the literature, two different methodologies are applied to identify the components of time series; thereby, Lee (1998), Blanchard, and Quan (1989) used a restricted VAR approach to determine the components. An alternative approach can be found in Tsay and Tiao (1990) who discussed several properties of multivariate non-stationary processes. Among other

---

[178] Note that the decomposition approach tries to identify a common stochastic trend of the three time series. Because the nominal capital is not cointegrated with share prices or dividends, the decomposition can be focused only on the two latter time series without any losses (see Tsay and Tiao, 1990, and Liu and Pan, 2003). If one uses three initial time series, the three components are later combined to a transitory and a permanent component.

[179] However, if all series were stationary, the resulting decomposition would lead to a single stationary component.

useful techniques, they introduced a canonical correlation analysis. This analysis can be applied to financial time series to decompose the transitory and long-run components of share prices, dividends, and earnings as shown by Liu and Pan (2003). Caused by its simplicity when dealing with panel data, I stick to the canonical correlation analysis. Thereby, the remainder is organized as follows. First, I highlight the calculation of canonical correlations. Second, after identifying which canonical variate is stationary, I determine the short and long-term components of share prices, dividends, and nominal capital. Thereafter, the transitory deviations are used to assess whether stocks are over- respectively undervalued.

5.8.3.2 Calculation of canonical correlations

Each element of z_{it} can be decomposed into a transitory and permanent component by solving the eigenvalues of the following matrix A_i. Thereby, this matrix consists of two coefficient matrixes that stem from a matrix regression of z_{it} on z_{it-1} and vice versa.

$$\mathbf{A_i} = \left(\sum_{t=2}^{T} \mathbf{z_{it} z'_{it}}\right)^{-1} \left(\sum_{t=2}^{T} \mathbf{z_{it} z'_{it-1}}\right) \left(\sum_{t=2}^{T} \mathbf{z_{it-1} z'_{it-1}}\right)^{-1} \left(\sum_{t=2}^{T} \mathbf{z_{it-1} z'_{it}}\right) \tag{5.22}$$

Note that the coefficient matrixes have the dimension 3×3 because z_{it} consists of three variables. What is the intuition behind this technique? If lagged values z_{it-1} possess the same partial impacts on current levels z_{it} than current values on lagged values, the matrix A_i will be equal to the identity matrix. And the 3×1 vector r_i that contains the ordered eigenvalues λ_{ij} will be the unity vector. This implies that one can determine three eigenvectors without imposing the normalization constraint. Let K_i be the partitioned matrix of the three eigenvectors k_{ij}.

$$\mathbf{K_i} = [\mathbf{k_{i1}}, \mathbf{k_{i2}}, \mathbf{k_{i3}}] \tag{5.23}$$

Then, K_i is equal to the identity matrix; hence, a decomposition of the initial time series is not possible. To see this, consider that the three time series cannot be expressed by a linear combination of their components as defined in equation 5.24.

$$\mathbf{z_{it}} = (\mathbf{K'_i})^{-1} \mathbf{K'_i z_{it}} = (\mathbf{K'_i})^{-1} \boldsymbol{\eta}_{\mathbf{it}} \tag{5.24}$$

Thus, in this extreme case, the canonical variates η_{it} are exactly the same as the initial time series z_{it}. Under normal conditions, however, the two coefficient matrixes deviate from one another; hence, a decomposition is possible. In addition, the matrix A_i is typically not symmetric which complicates the calculation of the associate eigenvectors k_{ij} because one has to impose the normalization constraint. Unfortunately, standard statistic programs have not the capability to calculate these eigenvectors.[180] Hence, I write my own STATA program to seek numerical solutions.[181] Thereby, one has to solve the non-linear system of equations.

$$\begin{aligned}&\mathbf{A_i k_{ij}} = \lambda_{ij}\mathbf{k_{ij}} \quad \forall i \text{ and } j = 1,2,3; \text{thereby } \lambda_{i1} \geq \lambda_{i2} \geq \lambda_{i3} \\ &\mathbf{k_i' k_i} = 1\end{aligned} \tag{5.25}$$

To carry out numerical methods in STATA that solve this non-linear system of equations, I have to use a `trick´. The first step is to define an artificial dependent variable labeled y = (1, 0, 0, 0). Now, I rearrange the system of equations so that the dependent variable y appears, for instance, on the right-hand site. Thus, the system one eigenvector k_{ij} (j=1,2,3) has the following shape. Note that the index 1,1 refers to the first element of the 3×1 dimensional vector.

$$\begin{aligned}&\left[\left(\mathbf{A_i} - \lambda_i \mathbf{I_{3\times3}}\right)\mathbf{k_i}\right]_{1,1} + 1 = 1 \\ &\left[\left(\mathbf{A_i} - \lambda_i \mathbf{I_{3\times3}}\right)\mathbf{k_i}\right]_{1,2} = 0 \\ &\left[\left(\mathbf{A_i} - \lambda_i \mathbf{I_{3\times3}}\right)\mathbf{k_i}\right]_{1,3} = 0 \\ &\mathbf{k_i' k_i} - 1 = 0\end{aligned} \tag{5.26}$$

Now, it is straightforward to derive the partitioned matrix K_i and to calculate the canonical variates η_{it}. After achieving the three canonical variates, one has to determine whether the respective variate is stationary or not by carrying out unit-root tests. By adding the weighted stationary canonical variates, the transitory component emerges, and the permanent component stems from the summed non-stationary weighted variates. To illustrate this procedure, it seems to be worthwhile to consider an example that is carried out step by step in the next section.

5.8.3.3 Components of the `Berliner Handelsgesellschaft´

To facilitate the understanding of the discussed econometric method, permanent and transitory components of the bank `Berliner Handelsgesellschaft´ are derived. First, the vector z_{it} is constructed which leads to the matrix A_i, after a simple calculus. Second, the eigenvalues are collected in vector r_i = (0.9997, 0.9267, 0.4029). Solving the non-linear system of equations, one gets the matrix K_i. Third, the canonical covariates follow from equation 5.24 and are tested using the Phillips-Perron and Dickey-Fuller procedure, which are presented in table 5.14. The first variate is non-stationary and hence represents the permanent component. In contrast, the second and third variate are stationary and add up to the transitory part of the time series.

[180] STATA can only determine characteristic vectors if the matrix A_i is symmetric.
[181] On request, a program (do-file) for STATA 8.0 is available from the author. The basic features of my program are also include in the statistical appendix.

Table 5.14: Testing the canonical covariates

This table contains the outcomes for the `Berliner Handelsgesellschaft´; thereby, I display the approximate p-values due to Mac Kinnon (1994) for the Phillips-Perron and Dickey-Fuller test applied to the three variates.

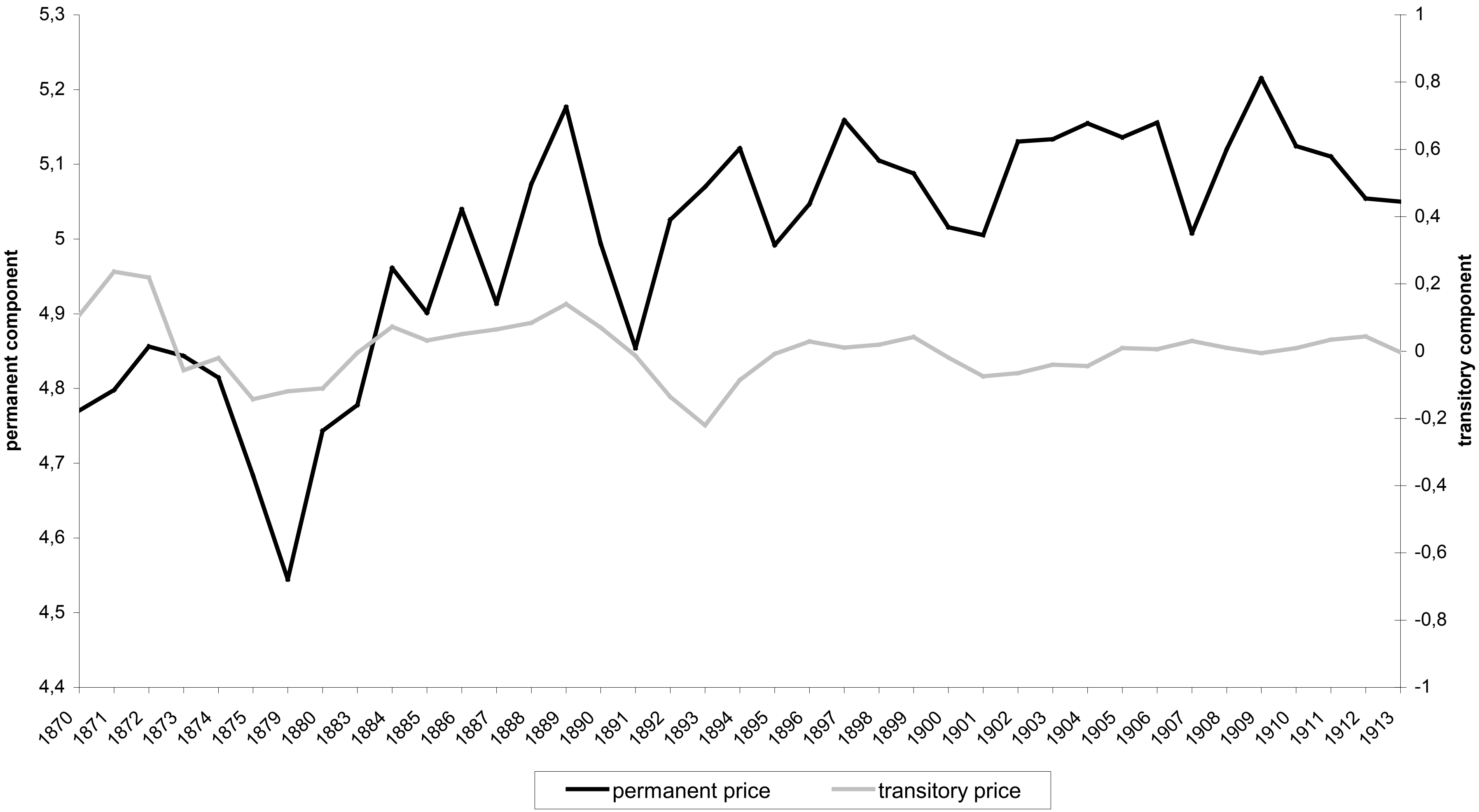
Canonical variates	Phillips-Perron	Dickey-Fuller
η_{i1}	0.0855	0.0754
η_{i2}	0.2729	0.4345
η_{i3}	0.1410	0.1539

One great advantage of this technique is to view the change of transitory and permanent components over time. For the purpose of illustration, figure 5.15 depicts the permanent and transitory part of the stock price during the period 1870 to 1913 of the `Berliner Handelsgesellschaft´. Apparently, the transitory component exhibits a stationary behavior and possesses negative and positive values, whereas the long-term component shows an upward tendency over this time period. This technique is normally used to identify different kinds of shocks. Accordingly, the permanent component of dividends represent fundamentals, whereas the transitory share price is due to short-term perturbations. Based on variance decompositions, one can assess whether variations in share prices stem from innovations in fundamentals. Liu and Pan (2003) used the decomposition for this `traditional´ approach to quantify the importance of non-fundamental factor for changes in share prices.

In contrast to this study, I favor another application that focuses on the transitory components of share prices and dividends. My consideration is close to the `spirit´ of the simple fad model (see West, 1988) in which the logarithmic price-dividend-ratio deviates from a permanent level through stationary perturbations. Due to taking logarithms, the difference between transitory share prices and transitory dividends represents a stationary perturbation or so called fad. Henceforth, a positive difference between transitory share prices and transitory dividends indicates overvaluation, whereas a negative difference shows that the stock is overvalued.

Figure 5.15: Permanent and transitory component of `Berliner Handelsgesellschaft's´ share price

The permanent part of the stock prices exhibit a non-stationary behavior, while the transitory component fluctuates around zero.



5.8.3.4 Transitory perturbations and the merger activity

5.8.3.4.1 The real valuation of the whole market and the merger wave

To make a statement whether the market as a whole is under- or overvalued, it is sufficient to decompose the components of equally weighted indexes. Figure 5.16 shows that for the whole sample, transitory share prices are below their permanent level during the period 1898 to 1911. Furthermore, dividends are higher than their permanent component might suggests. Hence, I state that even without estimating a cointegration relation and hence using restrictive assumptions, real undervaluation can be observed. To compare my results obtained by the cointegration analysis with the decomposition approach, figure 5.17 depict the cumulated transitory component of share prices and dividends from 1898 to 1911. In line with my former result, there is evidence of a considerable real undervaluation.

5.8.3.4.2 The real valuation in different industries and the merger activity

In addition, I focus on the real valuation in different industries, namely banking, mining, traffic, real estate, breweries, textile, machinery, metal, chemical industry and the category of other minor industries. This distinctions may give an impression whether the banking sector is undervalued compared to other lines of business. If a relative undervaluation were observed, the above-average merger activities in the banking industry could be explained.

The first analytical step is to decompose the equally weighted share prices and dividends for every industry into a permanent and transitory component. The subsequent step extents my former panel probit model to explain the decision to undertake a merger by inserting the transitory dividends and share prices. Note that for my discussion with regard to the impact of mergers on company characteristics, my former model is sufficient. The goal was to make the merger decision exogenous in my panel VAR and to construct forecasting errors that serve as micro-level shocks. However, this model is not appropriate to clarify the differences in the merger activity among industries. Hence, controlling for the real

Figure 5.16: Depicts the permanent and transitory component of share prices and dividends for all companies

The permanent parts exhibit a non-stationary behavior, while the transitory components fluctuate around zero. Note that I use an equally weighted portfolio.

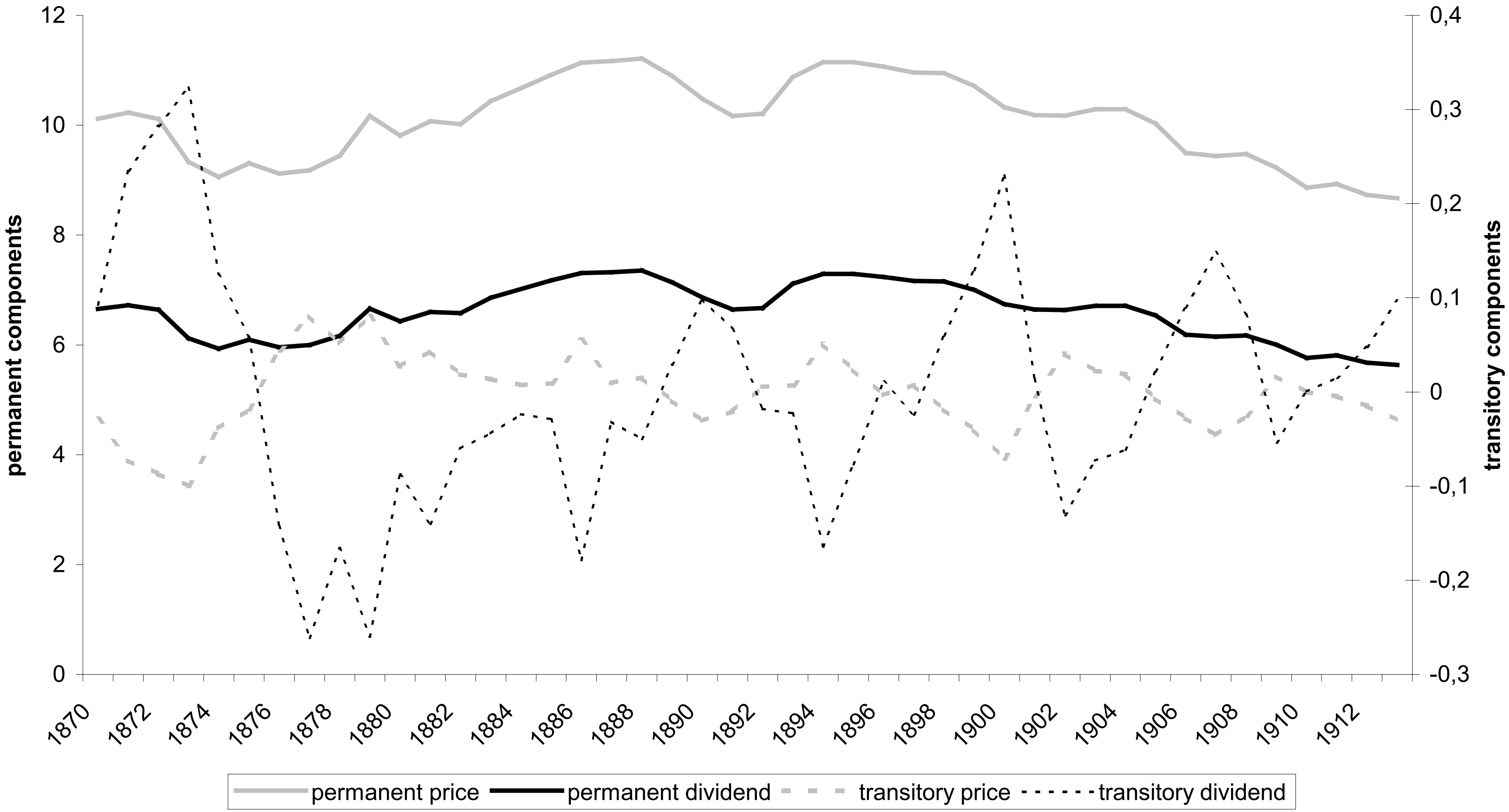
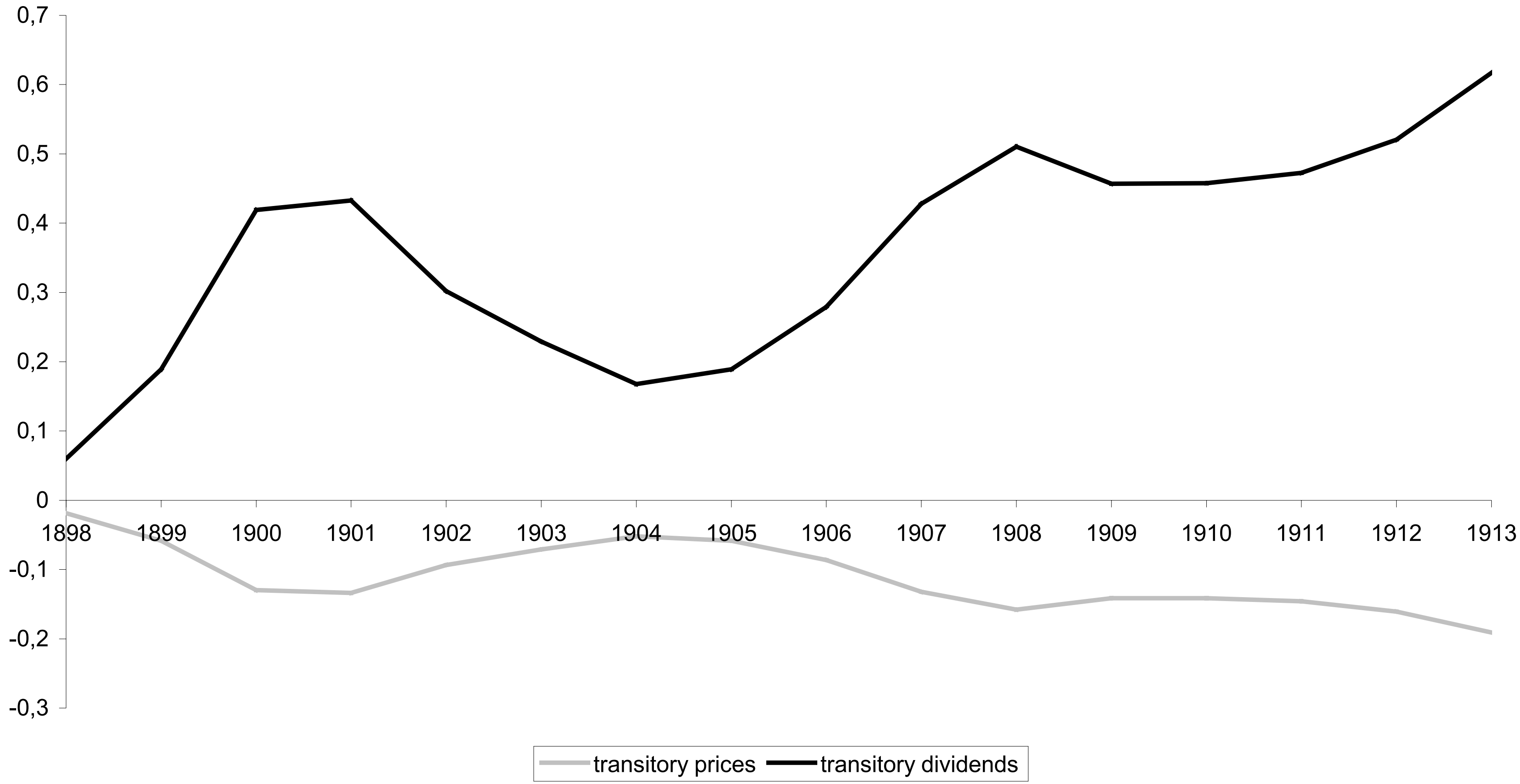


Figure 5.17: The transitory deviation of share prices and dividends is cumulated starting 1898 to 1913

The transitory components of share prices and dividends are added up over an increasing time interval; thereby, covering the period that exhibits the highest merger activity, namely 1898 to 1913.



valuation of companies within a specific line of business might clarify whether the real valuation explains the overwhelming merger activity among banks.

$$m_{it} = \alpha + \sum_{j=0}^{p} \boldsymbol{\beta}_{\mathbf{j}}' \boldsymbol{\Delta} \mathbf{z}_{\mathbf{it-j}} + \sum_{j=1}^{p} \gamma_j m_{it-j} + \sum_{j=1}^{p} \delta_j p_{kj}^t + \sum_{j=1}^{p} \lambda_j d_{kj}^t + u_i + \varepsilon_{it} \tag{5.27}$$

The former model is extended by the transitory components of share prices p^t and dividends d^t in the respective industry denoted k. Furthermore, the lag structure p is arbitrary and is determined by the Akaike criterion.

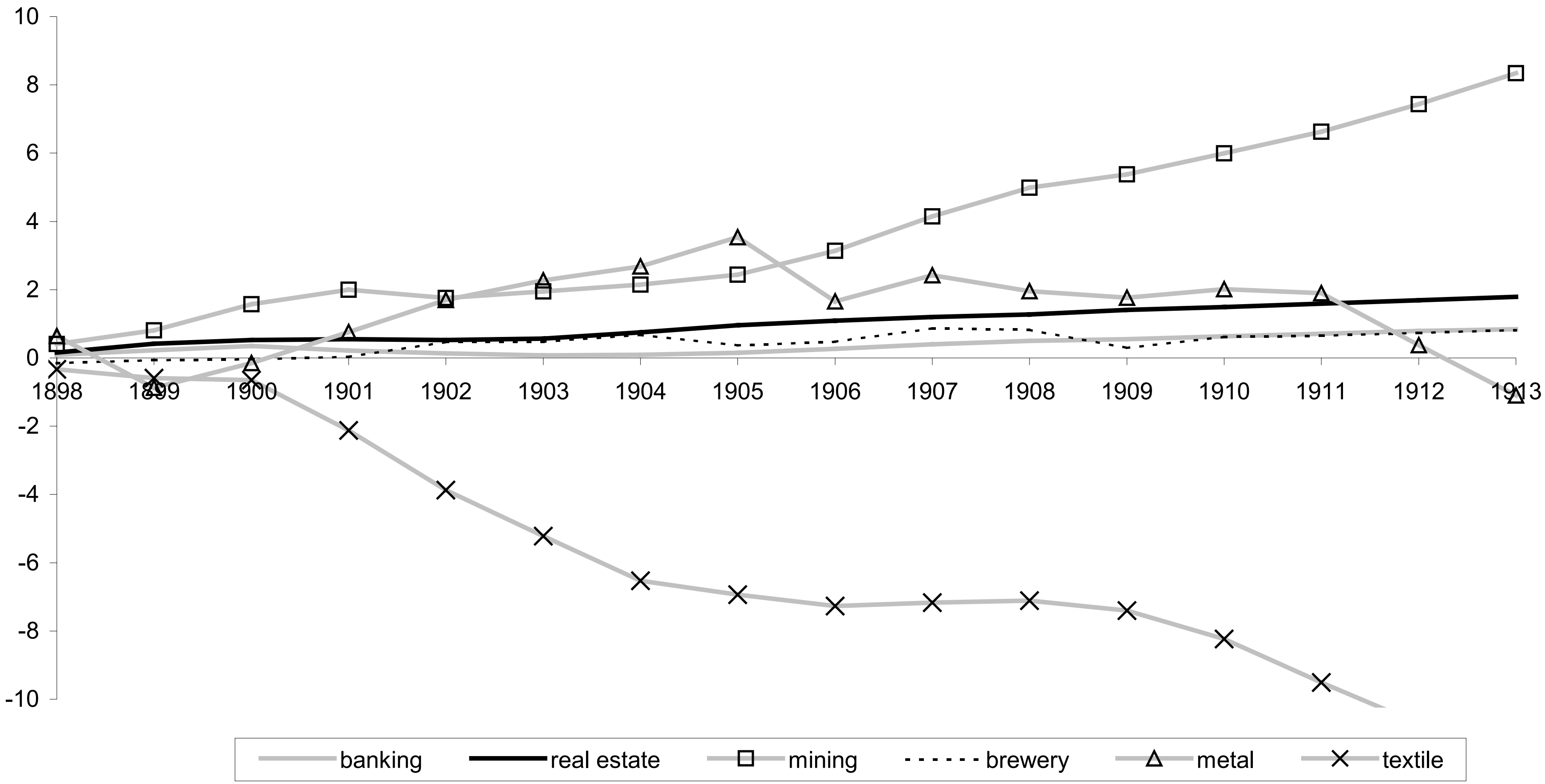
Unfortunately, my former model can hardly be improved. Inserting a dummy variable for the banking industry, however, is the only additional improvement that contributes to explaining the merger process. The dummy variable for banks is always significant (p-value: 0.000) and possesses a marginal impact of 3.52% on the probability for a merger. Because the panel probit model is very similar to my former specification and the results are also nearly identical, I skip the output table.

By inspiring figure 5.18, it becomes obvious that the real valuation of the banking industry is very close to the other lines of business. Hence, the results of the probit model are not surprising. To facilitate the illustration of real undervaluation during the period of high merger activity, namely 1898 to 1913, I depict the difference between transitory dividends and transitory share prices. The higher the difference the lower is the real valuation of the respective industry. Noteworthy, the textile industry is highly overvalued due to pronounced declines in dividend payments, whereas the mining companies exhibit a remarkable undervaluation. Nevertheless, all other industries are quite similar with regard to their real market valuation.

Based on these results, I conclude that the period between 1898 to 1913 exhibited a general real undervaluation of companies. Unfortunately, there is no evidence that the difference in the willingness to execute mergers among industries can be explained by the real market valuation of the respective line of business.

Figure 5.18: Real valuation in different lines of business

I calculated the difference between transitory dividends and transitory share prices; thereby, high values indicate short-term undervaluation. Furthermore, the annual differences are cumulated over an increasing time horizon until the whole period from 1898 to 1913 is covered.



5.9 The role of the exchange law 1896

The importance of the new exchange law established in the year 1896 for the creation of the German universal banking system is widely discussed in current literature. By far the most elaborate article on this issue is Fohlin's (2002) contribution. Without any doubts, there are some good qualitative arguments why the exchange law may affect the concentration process especially in the banking industry. For instance, the restrictions regarding forward dealings on exchanges imposed by the new law created a new profitable niche for large banks to conduct their own forward trading. However, from a quantitative point of view, evaluating whether the new law supported the extraordinary high merger activity from 1898 to 1913 is far from being straightforward. Note that the so called first merger wave 1898 to 1904 is a worldwide phenomenon. Thus far, there is descriptive evidence for the merger wave provided by Banerjee and Eckhard (2001) for the US case. My own investigation stresses that there are similar pattern for the German industry if one includes the banking industry. Besides the undisputable descriptive findings, the reasons for the first merger wave are less clear. For instance, Stigler (1950) called this concentration process a merging for monopoly and stressed the lack of regulations as ground for the merger wave. Because the first merger wave is a general phenomenon not only limited to Germany, the claim that a national change in the regulatory framework could trigger this immense movement is less convincing.

To quantify the impact of the new exchange law in the year 1896, I specify a dummy d_{law} variable that takes the value one for the period after the year 1896 and zero otherwise.

$$m_{it} = \alpha + \sum_{j=0}^{p} \boldsymbol{\beta}'_{\mathbf{j}} \Delta \mathbf{z}_{\mathbf{it-j}} + \sum_{j=1}^{p} \gamma_j m_{it-j} + \delta\, inf_t + \mu\, bank_i + \lambda d_{law} + u_i + \varepsilon_{it} \tag{5.28}$$

Table 5.15 presents the regression output together with the marginal effects. The period after the year 1896 could be characterized by a 3.70 percentage points higher probability that a company initiated a merger. My former results regarding the autocorrelation structure and the influence of higher inflation rates on the merger activity are still the same.

Table 5.15: Panel probit model with a dummy variable for the after 1896 period

To figure out whether the new exchange law of the year 1896 affected the probability of mergers, I estimate a simple panel probit model with random effects. Besides the coefficients, the marginal effects are presented, which facilitates the interpretation. The sufficient lag structure is chosen by the Akaike criterion.

	Coefficients	P-values	Marginal effects
Constant	-2.8250	0.000	-
m_{it-1}	0.7411	0.001	0.0565
m_{it-2}	0.8460	0.000	0.0711
Δn_{it}	1.0088	0.001	0.0366
$inflation_{it}$	0.0899	0.007	0.0033
$bank_i$	0.5691	0.000	0.0268
d_{law}	0.8020	0.000	0.0370
Number of observations	1437		
Wald Chi^2 statistic	103.47	0.000	

Despite, the highly significant coefficient of the dummy variable d_{law}, I would not claim that I find a clear evidence for the importance of the new exchange law. Based on my sample, it is impossible to distinguish between the pure time effect and the alleged influence of a regulatory change. To make a clear distinction, a cross-country study is needed in which some countries underwent a regulatory change, while others were in a stable institutional framework during the investigation period. Obviously, a cross-country study is out of the scope of my dissertation project – but an interesting subject for future research.

5.10 How important are mergers for the expansion of enterprises?

Despite the above-average merger activity among banks, the upsurge in nominal capital of banks is less superior compared to other industries. Figure 5.4 provide this impression. Thus, the question may arise whether mergers are mainly responsible for the expansion of enterprises during the first phase of globalization? Based on former models, I know that the change in nominal capital is not affected by current and past realizations of share prices as well as dividends. However, the model that predict the merger probability includes the change of nominal capital and confirms that an interrelation exists. Hence, I specify a panel model that explains the change in real nominal capital by taking into account mergers, macroeconomic conditions, and industry specific differences.

Table 5.16 summarizes the outcomes of different specifications; thereby, some models are estimated with fixed effects panel OLS, whereas others allow random effects and are estimated by GLS. F-tests indicate that all dummy variables for companies or points in time are abundant. However, random effects are also less convincing because the fraction of the total standard deviation of residuals that is due to individual effects approaches zero. A Breusch Pagan lagrangian multiplier test stresses the minor importance of individual effects. Thus, a simply system respectively panel OLS yields similar results, despite neglecting the panel structure.

To test for autocorrelation, I orient toward Wooldridge's (2002) two-step regression method; thereby, I estimate the following fixed effects panel regression with OLS.

$$\Delta n_{it} = \alpha_i + \beta_1 m_{it} + \beta_2 \text{inflation}_{\text{t}} + \beta_3 \text{gdp}_{\text{t}} + u_{it} \tag{5.29}$$

Thereafter, the residuals are estimated, lagged by one year, and inserted into the initial equation.

$$\Delta n_{it} = \alpha_i + \beta_1 m_{it} + \beta_2 \text{inflation}_{\text{t}} + \beta_3 \text{gdp}_{\text{t}} + \rho u_{it-1} + u_{it} \tag{5.30}$$

A standard t-test indicates whether the coefficient ρ is significantly different from zero. If ρ deviates considerably from zero, one can conclude that first order autocorrelation is present in the data. The last column of table 5.16 reports the outcome of regression 5.30 and underlines that autocorrelation does not exist. This two-step regression procedure assumes strict exogenity and by construction tests only for first order autocorrelation. Nevertheless, autocorrelation of higher order can be rejected by running the same procedure with more lags.[182]

[182] I skip these results.

Table 5.16: The expansion of nominal capital and the importance of mergers

I run a set of differently specified panel regressions to explain the change in real nominal capital; thereby, random as well as fixed effects representations are used.

	Random effects with GLS	Random effects with GLS	Fixed effects with OLS	Fixed effects with time dummies	Test for auto-correlation
Constant	0.0021	0.0075	0.0076	0.0578	0.0094
	(0.918)	(0.266)	(0.167)	(0.081)	(0.075)
m_{it}	0.0442	0.0438	0.0420	0.0624	0.0527
	(0.055)	(0.055)	(0.073)	(0.009)	(0.018)
$inflation_{it}$	-0.0199	-0.0199	-0.0198	-0.0105	-0.0225
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
gdp_{it}	0.0062	0.0062	0.0062	0.0116	0.0046
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Banking	-0.0006	-	-	-	-
	(0.980)				
Mining	0.0222	-	-	-	-
	(0.415)				
Traffic	0.0222	-	-	-	-
	(0.393)				
Real estate	-0.0235	-	-	-	-
	(0.472)				
Brewery	-0.0064	-	-	-	-
	(0.844)				
Textile	-0.0210	-	-	-	-
	(0.518)				
Machinery	-0.0040	-	-	-	-
	(0.901)				
Metal	0.0426	-	-	-	-
	(0.189)				
Chemical	0.0096	-	-	-	-
	(0.767)				
Others	0.0021	-	-	-	-
	(0.918)				
$Residual_{it-1}$	-	-	-	-	0.0387
					(0.126)
Observations	1456	1456	1456	1456	1415
Adjusted R^2	0.12	0.11	0.11	0.16	0.14

Arguing on the 10% level of significance, mergers and higher growth rates of NNP yield an increase in real nominal capital, whereas higher inflation rates prevent the expansion of enterprises. Furthermore, industry specific effects or company specific effects cannot be confirmed by my analysis.

5.11 Conclusion

Generally, one can draw the conclusion that macroeconomic conditions matter in this historical time period. More precisely, I can distinguish between a symmetric reaction of share prices and dividends triggered by changing economic growth, whereas fluctuations in inflation rates affect share prices stronger than dividends. Hence, in periods of increasing inflation, stocks are undervalued in real terms. Having in mind the relatively low volatility of stock markets in the period 1870 to 1913 in comparison to recent periods, macroeconomic factors possess a strong influence on stock prices and dividends. Furthermore, in line with the literature on the classical gold standard, I find that inflation and growth rates are hardly anticipatable. If a contemporary market participants use nominal interest rates, dividends, stock prices, and lagged values of the macroeconomic variables as information, they will not be able to make good one-step predictions.

From an econometric point of view, my study adds two new features to the analysis of the impact of macroeconomic factors on stock prices. First, by using a panel VAR approach, one obtains larger data sets that make the estimation of more complicated lag structures possible. Based on the exploitation of more information, the bootstrapping procedure yields relatively narrow confidence intervals. Second, I directly model the interrelation between share prices and dividends. This allows to analyze whether the response is an asymmetric or a symmetric one. Accordingly, I can assess whether the real valuation is affected by exogenous shocks.

Understanding the influence of macroeconomic shocks is essential to evaluate the long-term impact of mergers, which can be regarded as micro-level shocks. Embedding mergers as exogenous factors in my panel VAR faces the problem of endogenity with regard to the decision to execute a merger. Hence, I model the merger decision and uncover that share prices and dividends do not granger cause mergers. Based on this panel probit model with random effects, forecasts regarding the merger probability enable the anticipation of mergers. Obviously, errors in anticipating mergers are likely to cause market responses in share prices and dividends. Testing this hypothesis reveals that macroeconomic shocks predominate micro-level shocks. Mergers possess no long-term impact on share prices and dividends.

The concentration of the merger activity from 1898 to 1911 was a worldwide phenomenon and is called the first merger wave. Theoretical explanations for the emergence of merger waves cannot always be examined applying an empirical model. Nevertheless, the imagination due to Schenk (2001) that managers follow a minimax regret decision rule fits excellently to my empirical findings. Based on Palacios-Huerta's (2003) considerations that decisions should be serially correlated if `professionals play minimax´, finding a positive relation between successive mergers confirms that decision makers followed a minimax regret principal. Accordingly, I provide evidence for a merger wave in Germany that took place in the period 1898 to 1912 and reached its peak in 1906. Whether mergers were successful or not is not essential for Schenk's (2001) argumentation.

Despite uncovering a significant positive impact of the new exchange law established in the year 1896 on the probability of mergers, I cannot claim that the new legal environment was responsible for the merger mania in Germany. To achieve valid results, one has to construct a cross-country data set which goes clearly beyond the scope of my dissertation project.

I argue that the period from 1898 to 1911 exhibited real undervaluation of companies. This undervaluation is triggered by sudden upsurges in inflation rates. Macroeconomic fluctuations are in turn responsible for the pronounced time pattern of real stock price returns. To detect cointegration between share prices and dividends, the initial time series should be transformed to eliminate the influence of macro-level shocks. After mean-differencing share prices and dividends, the presence of cointegration can be confirmed by panel based tests as well as on the individual level. By applying a panel group dynamic ordinary least squares estimator, I measure the cointegration vector. Inspiring the plotted deviations from the long-run equilibrium characterized the period during which the merger wave reached its peak as a phase of real undervaluation. Even if I apply a completely different methodology that decompose time series into a transitory and permanent sequence, I achieve similar results. It should not be understated that the decomposition technique uses unmodified time series and does not require a cointegration relation. In addition, the decomposition is robust against structural breaks in a long-term equilibrium. Although the merger wave and a phase of real undervaluation coincided in the period 1898 to 1912, the valuation of companies can neither explain the `ignition´ of the merger wave nor the discrepancies with regard to the willingness to merge among different industries.

6. Concluding remarks

6.1 The success of mergers

6.1.1 Rejecting the merger paradox for the pre-World-War I period

Newspaper announcements regarding an imminent merger cause severe market responses in the short-run leading to considerable increases in market values of acquiring and target firms. Hence, based on my sample drawn in the year 1908, I can reject the presence of the merger paradox – shareholders of acquiring firms benefit from mergers. Accordingly, my empirical finding adds an additional piece to the picture whether mergers create shareholder value; thereby, I focus on the pre-World-War I period in Germany which was never studied before. Using daily returns to improve the statistical power of event-studies is completely new for the pre-1914 period. Most noteworthy, my study underlines the high degree of informational efficiency of stock exchanges because market reactions due to mergers are centered closely around the public release of information. This finding can also be confirmed using alternative approaches like transfer function models or panel GARCH models. Regardless which model is applied, they all point in the same direction: the market reacts fast.

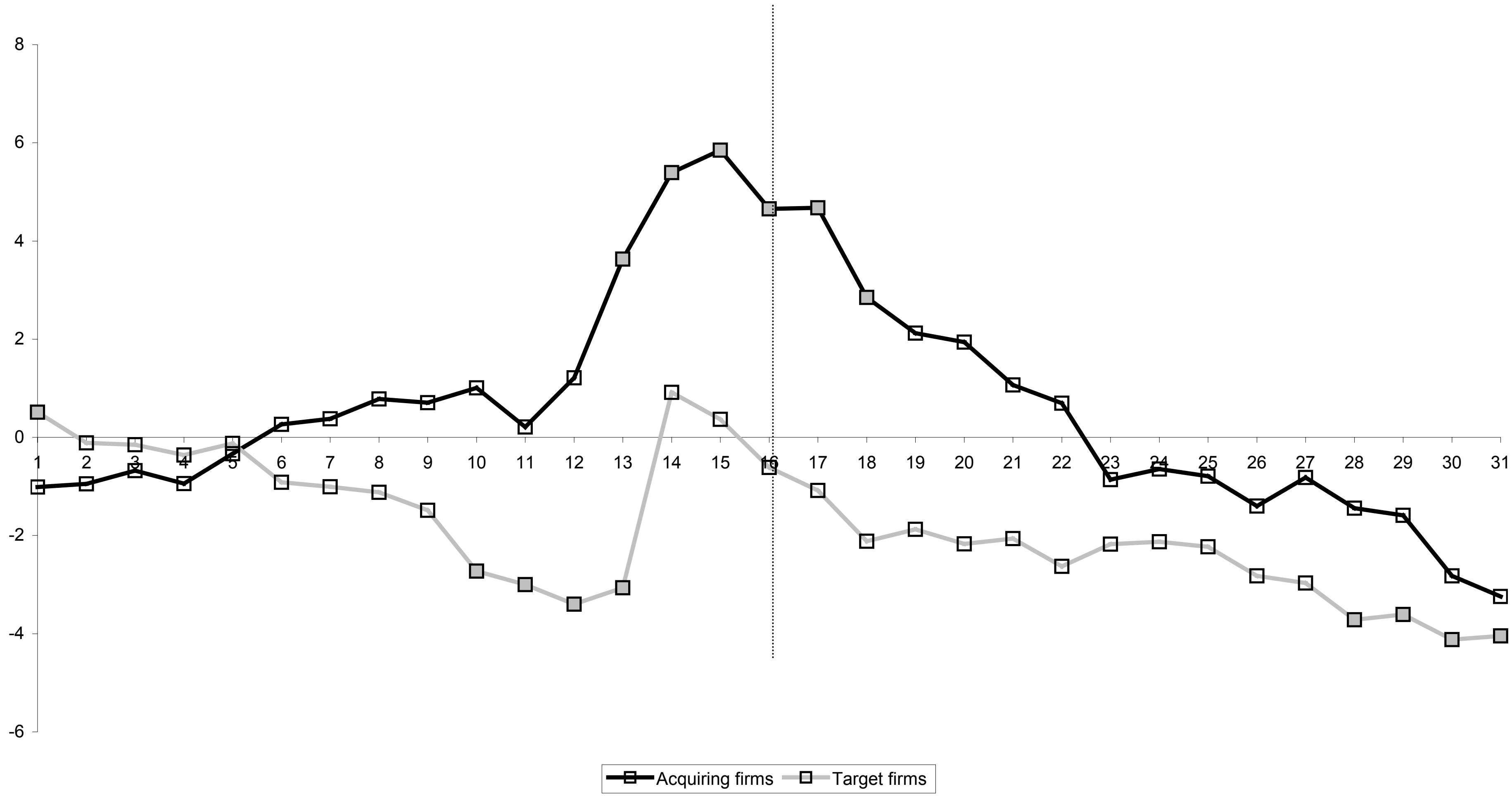
6.1.2 The emergence of the merger paradox in the year 2000

Dividing the sample of the year 2000 into acquiring and target firms uncovers that the merger paradox reappears even in a period of a high merger activity and a positive market environment.[183] Figure 6.1 plots the cumulated portfolio weighted abnormal return for the 34 acquiring companies against the same measure for the 27 target firms. Neither target firms nor acquiring companies gain from mergers – but the adaptation patterns differ. Three days before (t=13) the announcement untill two days afterwards (t=18) the market reacts positively in the

[183] Note that my study ends before the pronounced decline in share prices.

Figure 6.1: The re-emergence of the merger paradox

Figure 6.1 plots the aggregated cumulated abnormal return for increasing intervals starting at t=1 and ranging till t=31; thereby, I divide between acquiring and target firms. The vertical line indicates the announcement day.



case of acquiring firms and the cumulated abnormal returns show significant values. This short increase in market value is outweighed by the following slowdown that yields to a cumulated effect of –3.24% (p-value: 0.15) for the whole 31 days. Therefore, the market reacts ambiguous depending on the length of the time interval over which I aggregate – but the long-term effect is negative and insignificant. In contrast to the massive upsurge in market values of target firms in the year 1908, the market reacts negatively in the year 2000. Six days (t=10) to three days before (t=13) the public announcement, a significant negative cumulated effect is observed. Moreover, the long-term effect over the whole 31 days is –4.05% (p-value: 0.07), and, hence, the shareholders of target firms loose on average. This empirical finding can be seen as the re-emergence of the merger paradox because all involved companies show decreasing market values over the whole event period. For Germany during the period 1971 to 1985, Bühner (1991) uncovered that acquirers lost on average six per cent after the merger. Thus, he confirmed the merger paradox. However, my finding that targets exhibit a pronounced decline in share prices was seldom found by other studies.[184]

Accordingly, my research for the pre-World-War I period underlines that in Germany successful mergers indicated by increasing market values of target and acquiring companies were possible. Section 6.6 provides some descriptive and narrative evidence why this empirical finding is plausible and in line with theoretical argumentations.

6.2 Which shareholders gain from mergers?

6.2.1 Insider trading in the year 1908

Going 100 years back in history, natural firm behavior without regulatory restrictions is observable; thereby, firms could decide to disclose price-sensitive information voluntarily. Event-studies and cross-sectional models confirm that hiding information is chargeable to outsiders. Ruling out the possibility that the way of disclosure influences the total gains from mergers, one can concentrate on the distributive effect. According to protecting outsider, a state intervention that forces firms to release information should be considered; particularly, a voluntary self-regulation cannot be supported by logit models. Henceforth, this historical experiment stresses that ad-hoc-publication requirements or other retaliations, like a negative public opinion regarding the misbehavior, ensure the protection of outsiders. Most noteworthy, the media in 1908 did not criticize that acquirers started buying shares of targets

[184] Cosh and Guest (2001) found that targets' share prices decline on average by 18% if the transaction can be characterized as hostile takeover. Even if the merger is a `friendly offer´, they could not detect significant abnormal returns for target firms. They studied mergers in the U.K. during the period 1985 to 1996.

prior to official announcements – though newspapers spread rumors about imminent transactions.

6.2.2 Irrational speculation in the year 2000

The adaptation process of share prices in the presence of newly available information differs from the clear run-ups and steady increases in market values observed in the historical sample. Generally, a strong upsurge in share prices shortly before the newspaper announcement is followed by a pronounced fall in market values. Regardless which group of companies is considered, this pattern remains nearly unaffected. Correspondingly, this adaptation pattern could stem from following irrational trading rules like `buy on rumors and sell on facts´. This empirical finding supports the effectiveness of insider regulation during the last 92 years.

6.3 Interpreting the high informationally efficiency in 1908

Not only the event-studies but also GARCH models point out that the market in 1908 is highly informationally efficient. Besides the remarkable actuality of the `Berliner Börsenzeitung´ as shown by several case studies, insider trading may be mainly responsible for the tremendous speed with which new information is reflected in market prices. Accordingly, the benefit of informational efficiency comes with a loss, namely insider trading. Thus far, little historical research was conducted to quantify the number of insiders versus outsiders during the pre-1914 period in Germany. Since the new exchange law established in 1896, the nine leading banks in Berlin gained a larger role in trading. This qualitative statement supports the view that insiders dominated the trading during the pre-1914 period in Germany. For the U.S., Warshow (1924) pointed out that smaller shareholders, the `typical´ outsiders, were relatively unimportant.[185]

6.4 The long-term impact of mergers

Mergers characterized as micro-level shocks possess a significantly negative impact on share prices; thereby, this direct impact also affects dividend streams with a time lag. After taking into account unexpected macroeconomic shocks, the effect of mergers disappears; hence, macro-level shock predominate in the period 1870 to 1914. Nevertheless, additional insights are gained by extending my short-term evidence based on event-studies: macroeconomic

[185] In table II, he presented some figures for the U.S. corporate ownership.

surprises severely affect share prices and dividends. Obviously, focusing on market responses within 31 days, one fails to measure macroeconomic shocks.

6.5 Merger waves and periods of real over- or undervaluation

As far as I know, chapter five provides the first empirical evidence for a merger wave in Germany centered around 1906. Executed mergers one or two years ago increase the likelihood for subsequent transactions significantly. Furthermore, mergers are more likely in periods exhibiting high inflation rates – but past and present share prices and dividends have no partial impact regardless which lag structure is permitted. Based on impulse response functions, an asymmetric response – triggered by shocks in inflation rates – of share prices and dividends can be observed. Hence, an unanticipated upsurge in inflation causes a phase of real undervaluation of equity. More formal models like hidden cointegration and the decomposition of time series yield similar outcomes. Consequently, I state that the first merger wave coincided with a period of real undervaluation of companies. A recent study for the period from 1978 to 2000 conducted for the United States of America by Dong et al. (2003) showed that mergers occur if markets are overvalued. Hence, my empirical finding for the first phase of globalization in Germany contradicts today's empirical evidence.

6.6 What makes mergers different: Comparing both phases of globalization

Besides the hard facts justified by empirical models, adding some narrative evidence told by my case studies helps to obtain a clearer picture. During my investigation period, cross-boarder mergers did not occur – albeit very common nowadays. Furthermore, acquirers were relatively large compared to their target firms making an acquisition easer to finance and facilitating the integration of both entities. The velocity with which mergers were legally executed in the pre-1914 period is remarkable. As shown by the presented case studies, the announcement of a merger is followed by the approval of an extraordinary shareholder gathering within one month usually. Nearly all mergers achieved the necessary majority in shareholder gatherings, and hostile takeovers were very uncommon. Although the leading companies included in my long-term study were very active to conduct mergers, mergers among these 35 largest companies did not occur. Accordingly, mergers are to some extent different in both phases of globalization – but maybe learning from historical evidence could help to make today's merger as successful as 100 years ago. A possible recommendation could state: acquire only smaller companies in your line of business and be friendly!

6.7 The inflation illusion hypothesis and the pre-World-War I period

In chapter five, my empirical investigation supports the inflation illusion hypothesis for the period 1870 to 1914 in Germany – but how can one interpret this finding. Since 1873, Germany joined the gold standard; thus, one can argue that by introducing an effective commitment to avoid inflation an inflation hedge provided by stocks was not needed. Unfortunately, this explanation has a blemish in that the inflation illusion can also be observed in later period.[186]

Inspiring figure 5.1, the cyclical movement in inflation rates characterized by alternating periods of inflation and deflation is apparent. Accordingly, I argue that during the pre-World-War I period trend inflation cannot be observed; thereby, the average annual inflation rate is lower than 1%. Considering a long-term investor, namely a bank or a strategic investor, the question arises whether market participants should worry about inflation in the long-run. Due to the overwhelming importance of large investors with strategic interest, nominal share prices should not reflect inflation and inflation illusion is likely.

In addition, time series of nominal interest rates show that during the pre-World-War I period the development was almost stable over time; thus, inflation rates did not influence nominal interest rates considerably. Putting this argument in other words, it states that by observing nominal interest rates market participants were not able to improve their forecasts regarding future inflation rates. This finding is supported by my analysis on the predictability of macroeconomic variables. Hence, stable nominal interest rates suggest that inflation was not a major concern during the pre-1914 period. Figure 6.2 depicts the inflation rate based on Hoffmann (1965), the private discount rate, the discount rate set by the \`Reichsbank´ denoted RB discount, and the average yield of loan stocks.[187]

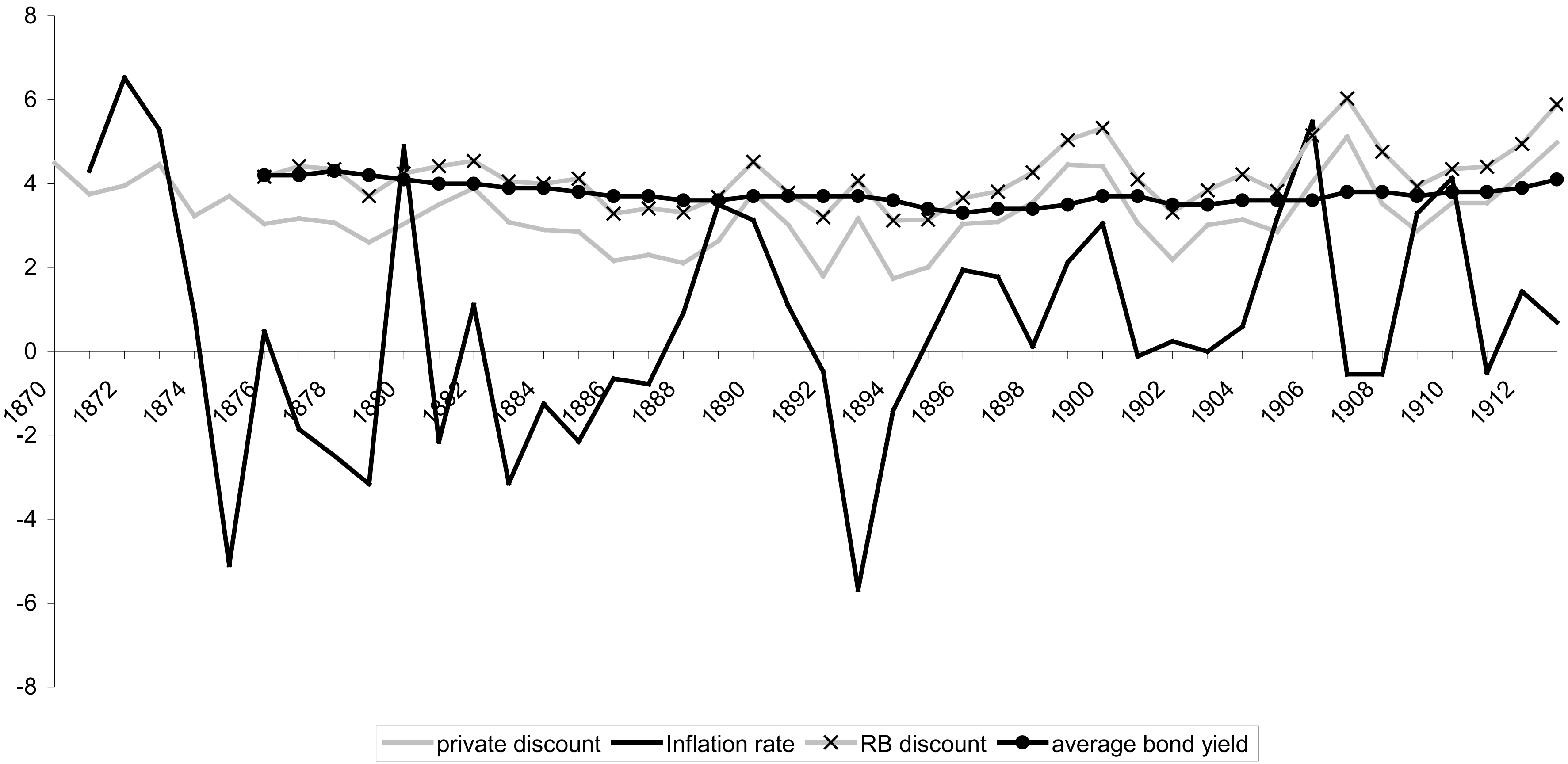
---

[186] Madsen (2002) tested the inflation illusion hypothesis with a panel data set consisting of the OECD countries; thereby, he focused on two periods, namely the post-World-War II and the Great Depression. The estimates suggested that share markets suffered from inflation illusion.

[187] Interest rates after 1876 can be found in \`Deutsches Geld- und Bankwesen in Zahlen 1876-1975 (1976)´, and Graba's (1992) data appendix contained private discount rates for the period before 1876.

Figure 6.2: Nominal interest rates and inflation

Figure 6.2 plots inflation rates and three different nominal interest rates for the period from 1870 to 1876; thereby, Hoffmann (1965), Grabas (1992), and the data collection entitled `Deutsches Geld- und Bankwesen in Zahlen 1876-1975 (1976)´ are used as sources.



6.8 Was the effort worth it?

Reading daily newspapers for several months, collecting about 6550 daily returns for the sample of the year 1908, 4941 daily returns of individual stocks and several thousands of daily observations of the market index, DAX30, for the year 2000 as well as 4620 observations of share prices, dividends, and nominal capital for the long-term study, one should wonder whether the effort was worth it.

Definitely, many new results were obtained by my research like the rejection of the merger paradox and its re-emergence and the empirical evidence for the first merger wave in Germany. Henceforth, I contribute in many respects to the existing body of literature in economic history.

As mentioned in the preface, empirical results should always be handled with caution and I hope that I also stressed the problems inherent with data sets and empirical methods. In empirical research, the way to achieve a result is much more interesting than the result itself. Accordingly, the proposed alternative approaches and thorough debates about the pros and cons of every method supplement the literature and might be useful for other applications.

Although considerable insights regarding the success of mergers, insider trading, and the influence of regulation contribute to the understanding of firm behavior in different periods in Germany, there is – as always – still enough space for future research. Especially, cross-country studies might be promising to determine the role of the exchange law 1896 on merger activities reliably. Furthermore, a consumption based valuation of asset returns could improve our understanding of the pre-1914 period because the amazing importance of macroeconomic factors point into this direction.

The fourth chapter stresses some limitations of the short-term analysis and alternative models; thereby, nearly every section can be extended to an independent paper. For instance, my GARCH approach could be also applied to the sample of the year 2000 respectively to individual stocks. Furthermore, the way of disclosure might also affect the volatility pattern of stock returns. Modeling trading patterns can be the starting point for a quantitative paper on the market microstructure of the Berlin stock exchange around 1900. Besides providing quantitative evidence regarding early financial market in Germany, qualitative research is still a promising field for future research. Although Berghoff (2002)[188] described precisely the

[188] His article is part of an edited volume that contains additional contributions due to Buchheim, and Mieck who also focused on the pre-1914 period.

history of the Berlin stock market, many aspects seem to be still attractive for future research like the role of forward dealings, speculation, and state interventions.[189]

Moreover, one should keep in mind that my thesis is one way to discuss the success of mergers, besides already mentioned alternatives,[190] accounting-based measures to evaluate the performance after acquisitions are quite common. If one concentrates like Ukaegbu (1987) on the period from 1969 to 1984, an analysis that relies mainly on balance-sheets is feasible. In contrast, for my investigation period from 1870 to 1914, such an analysis is nearly impossible due to lacking or time-varying accounting standards.

Accordingly, every study is by nature limited; however, I provided several new empirical findings, explored new data sources, and developed some `crazy´ ideas to analyze my data sets.

Finally, one should cite the famous words due to McCloskey (1994) who stated: "(...) empirical research in economics is unbelievable, uninteresting or both." I deeply hope that his impression does not apply to my research.

---

[189] To be more precise, Berghoff (2002) discussed the forward dealings (`Zeitgeschäfte´) that had foreign bonds as underlyings. He also noted that due to a tremendous increase in speculation, regulations were introduced in 1840 to forbid these `speculative dealings´ - albeit private market makers could overcome this restriction. He also referred to forward dealings based on railroad companies in the 1840s and the first `railroad mania´ around 1844 which led to an anti-capitalistic state intervention with regard to speculation.

[190] Tilly (1982) proposed survivorship as measure of success and Jarrell et al. (1988) highlighted the redistribution theory, which is based on a stakeholder concept.

7. Mathematical and statistical appendix

7.1 Mathematical appendix

7.1.1 Technical note for equation (2.6)

Define matrix **A** as n×L dimensional matrix that contains all stocks i and for each point in time l all observed daily returns. I also define the unit vector **1** as L×1 dimensional vector. Thus, I can write:

$$L^{-1} \cdot \mathbf{A1} = L^{-1} \cdot \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1L} \\ R_{21} & R_{22} & \dots & R_{2L} \\ \vdots & \vdots & \vdots & \vdots \\ R_{n1} & R_{n2} & \dots & R_{nL} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = L^{-1} \cdot \sum_{l=1}^{L} \mathbf{R}_l$$

7.1.2 Proof of proposition 1

Taking the conditional expectation with respect to the data matrix **A** of equation (2.9) determines the conditional mean vector.

$$E\left(\hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^{*} \middle| \mathbf{A}\right) = E\left(\mathbf{R}_{\mathbf{t}} \middle| \mathbf{A}\right) - \boldsymbol{\mu} = \mathbf{0}$$

This follows from equation (2.4) as long as the error term e_l possesses its desired properties, especially, I can assume the following

$$E(\mathbf{e}_{\mathbf{l}} | \mathbf{A}) = \mathbf{0}$$

…and the MM estimator of the mean vector is unbiased.

$$E\left(\hat{\boldsymbol{\mu}} \middle| \mathbf{A}\right) = \boldsymbol{\mu}$$

Afterwards, I am interested in the conditional covariance matrix $\mathbf{V_t}$ of the estimated abnormal returns. Because the conditional mean vector is equal to the zero vector as shown above and by using expression (2.9), the covariance matrix is…

$$\text{(I)}\ \mathbf{V}_{\mathbf{t}} = E\left(\hat{\boldsymbol{\varepsilon}}_t^{*} (\hat{\boldsymbol{\varepsilon}}_t^{*})' \middle| \mathbf{A}\right) = E\left(\left(\mathbf{R}_{\mathbf{t}} - \hat{\boldsymbol{\mu}}\right)\left(\mathbf{R}_{\mathbf{t}} - \hat{\boldsymbol{\mu}}\right)' \middle| \mathbf{A}\right)$$

The CMR model tells us that one can express the vector of observed returns by using the mean vector and error term as shown in (2.4). Note that the error term e_l can be estimated as residual using the observations of the estimation period

(II)

$$\overset{(I)}{\Rightarrow} E\left(\left(\mathbf{e}_{\mathbf{l}} + \boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\right)\left(\mathbf{e}_{\mathbf{l}}' + \boldsymbol{\mu}' - \hat{\boldsymbol{\mu}}'\right) \middle| \mathbf{A}\right) = E\left(\mathbf{e}_{\mathbf{l}}\mathbf{e}_{\mathbf{l}}' + \mathbf{e}_{\mathbf{l}}\boldsymbol{\mu}' - \mathbf{e}_{\mathbf{l}}\hat{\boldsymbol{\mu}}' + \boldsymbol{\mu}\mathbf{e}_{\mathbf{l}}' + \boldsymbol{\mu}\boldsymbol{\mu}' - \boldsymbol{\mu}\hat{\boldsymbol{\mu}}' - \hat{\boldsymbol{\mu}}\mathbf{e}_{\mathbf{l}}' - \hat{\boldsymbol{\mu}}\boldsymbol{\mu}' + \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}' \middle| \mathbf{A}\right)$$

Now I discuss this expression term by term.

$$E\left(\mathbf{e}_1\mathbf{e}_1' \middle| \mathbf{A}\right) = E\left(\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}\begin{pmatrix} e_1 & e_2 & \dots & e_n \end{pmatrix}\middle| \mathbf{A}\right) = E\left(\begin{pmatrix} e_1e_1 & e_1e_2 & \dots & e_1e_n \\ e_2e_1 & e_2e_2 & \dots & e_2e_n \\ \vdots & \vdots & \vdots & \vdots \\ e_ne_1 & e_ne_2 & \dots & e_ne_n \end{pmatrix}\middle| \mathbf{A}\right) =$$

$$\begin{pmatrix} E(e_1e_1|\mathbf{A}) & E(e_1e_2|\mathbf{A}) & \dots & E(e_1e_n|\mathbf{A}) \\ E(e_2e_1|\mathbf{A}) & E(e_2e_2|\mathbf{A}) & \dots & E(e_2e_n|\mathbf{A}) \\ \vdots & \vdots & \vdots & \vdots \\ E(e_ne_1|\mathbf{A}) & E(e_ne_2|\mathbf{A}) & \dots & E(e_ne_n|\mathbf{A}) \end{pmatrix}$$

Note that error terms are supposed to be uncorrelated between cross sectional observations and the conditional mean of the error term vector is zero, which stems from the assumption introduced above. Hence, one allows that the variance of the error term vary across different stocks; however, the variance is constant over time. Thus, it follows.

$$\text{(III)}\ E\left(\mathbf{e}_1\mathbf{e}_1' \middle| \mathbf{A}\right) = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} = \mathbf{I}_{n\times n}\boldsymbol{\sigma}_e^2$$

Thereby, $\mathbf{I}_{n\times n}$ denotes the n×n dimensional identity matrix and $\boldsymbol{\sigma}_\mathbf{e}^\mathbf{2}$ represents the n×1 dimensional error variance vector.

To proceed, one has to show that the following equality holds.

$$E(\hat{\boldsymbol{\mu}}\boldsymbol{\mu}') = E(\boldsymbol{\mu}\hat{\boldsymbol{\mu}}')$$

Note, this also holds for conditional expectations (LIE: Law of iterated expectations).

The left hand side of this equality can be written as…

$$E\left(\begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \vdots \\ \hat{\mu}_n \end{pmatrix}\begin{pmatrix} \mu_1 & \mu_2 & \dots & \mu_n \end{pmatrix}\right) = \begin{pmatrix} \mu_1E\hat{\mu}_1 & \mu_2E\hat{\mu}_1 & \dots & \mu_nE\hat{\mu}_1 \\ \mu_1E\hat{\mu}_2 & \mu_2E\hat{\mu}_2 & \dots & \mu_nE\hat{\mu}_2 \\ \vdots & \vdots & \vdots & \vdots \\ \mu_1E\hat{\mu}_n & \mu_2E\hat{\mu}_n & \dots & \mu_nE\hat{\mu}_n \end{pmatrix}$$

If the estimator of the mean vector is unbiased, I can conclude

$$E(\hat{\boldsymbol{\mu}}) = \boldsymbol{\mu}$$

Using this finite sample property of the MM estimator, it appears to be obvious that the quadratic n×n matrix is also a symmetric matrix.

Note:

$$(\hat{\boldsymbol{\mu}}\boldsymbol{\mu}')' = \boldsymbol{\mu}\hat{\boldsymbol{\mu}}'$$

The transpose of this symmetric matrix is in turn the same matrix as before transposing the matrix. Thus, I can show that the equality is fulfilled.

Furthermore, I can infer that…

$$E\left(\begin{pmatrix}\hat{\mu}_1\\ \hat{\mu}_2\\ \vdots\\ \hat{\mu}_n\end{pmatrix}\begin{pmatrix}\mu_1 & \mu_2 & \dots & \mu_n\end{pmatrix}\right)=\begin{pmatrix}\mu_1 E\hat{\mu}_1 & \mu_2 E\hat{\mu}_1 & \dots & \mu_n E\hat{\mu}_1\\ \mu_1 E\hat{\mu}_2 & \mu_2 E\hat{\mu}_2 & \dots & \mu_n E\hat{\mu}_2\\ \vdots & \vdots & \vdots & \vdots\\ \mu_1 E\hat{\mu}_n & \mu_2 E\hat{\mu}_n & \dots & \mu_n E\hat{\mu}_n\end{pmatrix}=\begin{pmatrix}\mu_1\mu_1 & \mu_2\mu_1 & \dots & \mu_n\mu_1\\ \mu_1\mu_2 & \mu_2\mu_2 & \dots & \mu_n\mu_2\\ \vdots & \vdots & \vdots & \vdots\\ \mu_1\mu_n & \mu_2\mu_n & \dots & \mu_n\mu_n\end{pmatrix}=E(\boldsymbol{\mu}\boldsymbol{\mu}')$$

Hence, the following extended equality holds…

(IV) $E(\hat{\boldsymbol{\mu}}\boldsymbol{\mu}')=E(\boldsymbol{\mu}\hat{\boldsymbol{\mu}}')=E(\boldsymbol{\mu}\boldsymbol{\mu}')$

Using equality (IV) and the result (III), expression (II) simplifies to…

(V) $\overset{(II)}{\Rightarrow}\mathbf{I}_{n\times n}\boldsymbol{\sigma}_e^2+E(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}'|\mathbf{A})-E(\boldsymbol{\mu}\boldsymbol{\mu}'|\mathbf{A})$

It is straightforward to see that…

$$E(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}'|\mathbf{A})-E(\boldsymbol{\mu}\boldsymbol{\mu}'|\mathbf{A})=\begin{pmatrix}E(\hat{\mu}_1\hat{\mu}_1|\mathbf{A})-\mu_1\mu_1 & E(\hat{\mu}_1\hat{\mu}_2|\mathbf{A})-\mu_1\mu_2 & \dots & E(\hat{\mu}_1\hat{\mu}_n|\mathbf{A})-\mu_1\mu_n\\ E(\hat{\mu}_2\hat{\mu}_1|\mathbf{A})-\mu_2\mu_1 & E(\hat{\mu}_2\hat{\mu}_2|\mathbf{A})-\mu_2\mu_2 & \dots & E(\hat{\mu}_2\hat{\mu}_n|\mathbf{A})-\mu_2\mu_n\\ \vdots & \vdots & \vdots & \vdots\\ E(\hat{\mu}_n\hat{\mu}_1|\mathbf{A})-\mu_n\mu_1 & E(\hat{\mu}_n\hat{\mu}_2|\mathbf{A})-\mu_n\mu_2 & \dots & E(\hat{\mu}_n\hat{\mu}_n|\mathbf{A})-\mu_n\mu_n\end{pmatrix}$$

Obviously, this expression denotes the variance covariance matrix of the estimator of the mean vector. One imposes the assumption that the variance of the estimated means can vary across stocks i. Moreover, lets suppose that the estimated means are uncorrelated between different stocks i, which seems to be very plausible, except for stocks belonging to the same line of business.

Now I can simplify the variance matrix in the following manner:

$$E(\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}'|\mathbf{A})-E(\boldsymbol{\mu}\boldsymbol{\mu}'|\mathbf{A})=\begin{pmatrix}Var\hat{\mu}_1 & 0 & \dots & 0\\ 0 & Var\hat{\mu}_2 & \dots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \dots & Var\hat{\mu}_n\end{pmatrix}=\mathbf{I}_{n\times n}Var(\hat{\boldsymbol{\mu}})$$

Accordingly, the conditional covariance matrix is now derived by joining together my former results.

(VI) $\overset{(V)}{\Rightarrow}Var(\hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^*|\mathbf{A})=\mathbf{I}_{n\times n}\boldsymbol{\sigma}_e^2+\mathbf{I}_{n\times n}Var(\hat{\boldsymbol{\mu}})$

q.e.d.

7.1.3 Proof of proposition 2

I calculate the conditional mean, given that one observes the data matrix **A,** of equation (2.11); thereby, the implications of proposition 1 are embedded.

$$E\left(\hat{\mathbf{C}}(\tau_m;\tau_n)\middle|\mathbf{A}\right) \equiv \sum_{t=\tau_m}^{\tau_n} E\left(\hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^{*}\middle|\mathbf{A}\right) = \mathbf{0}$$

Now, I turn to the conditional covariance matrix $\mathbf{V_C}$. Recall that the assumption that the variance of an abnormal return of stock i is constant over time is essential – but different variances across securities are permitted.

$$Var\left(\hat{\mathbf{C}}(\tau_m;\tau_n)\middle|\mathbf{A}\right) \equiv \sum_{t=\tau_m}^{\tau_n} Var\left(\hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^{*}\middle|\mathbf{A}\right) = T_1 \cdot \left(\mathbf{I}_{n\times n}\boldsymbol{\sigma}_e^2 + \mathbf{I}_{n\times n}Var(\hat{\boldsymbol{\mu}})\right)$$

Thereby, T_1 denotes the length of the interval over which abnormal returns are cumulated. Note that T_1 can differ from the length of the whole event period T. This formula also gives insight regarding the proper length of event windows. Correspondingly, the larger the event window, over which one wants to aggregate, the larger the variance of cumulated abnormal returns (CAR). If I want to evaluate long-term wealth effects of mergers using cumulated abnormal returns, this fact limits the possibilities to detect a deviation from the null hypothesis, caused by the high variance.

Using the property of normal distributions, the sum of normally distributed abnormal returns is in turn normally distributed. Because I have to estimate the variance and mean by method of moments (MM), the standardized CAR is t-distributed and for $T_1>30$ approximately standard normally distributed.

7.1.4 Proof of Proposition 3

A formal proof is not necessary because I use only the definition of the arithmetic middle and the calculation of the variance of uncorrelated abnormal returns at time t. However, I try to illustrate the formulae.

$$\bar{\varepsilon}_t^{\,*} = n^{-1} \cdot \left(\hat{\boldsymbol{\varepsilon}}_{\mathbf{t}}^{*}\right)' \mathbf{1} = n^{-1} \cdot \left(\hat{\varepsilon}_{1t}^{*} \quad \hat{\varepsilon}_{2t}^{*} \quad \dots \quad \hat{\varepsilon}_{nt}^{*}\right) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = n^{-1} \cdot \sum_{i=1}^{n} \hat{\varepsilon}_{it}^{*}$$

$$Var\left(\bar{\varepsilon}_t^{\,*}\right) = n^{-2} \cdot tr\left[\mathbf{I}_{n\times n}\boldsymbol{\sigma}_{\mathbf{e}}^{\mathbf{2}} + \mathbf{I}_{n\times n}Var(\hat{\boldsymbol{\mu}})\right] = n^{-2} \cdot tr\begin{bmatrix} \sigma_{e1}^2 + Var(\hat{\mu}_1) & 0 & \dots & 0 \\ 0 & \sigma_{e2}^2 + Var(\hat{\mu}_2) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_{en}^2 + Var(\hat{\mu}_n) \end{bmatrix}$$

Thereby, \`tr´ denotes the trace of the covariance matrix. The trace of a matrix is the sum of its first diagonal elements which simplifies this expression.

$$Var\left(\overline{\varepsilon}_t^*\right) = n^{-2} \cdot \sum_{i=1}^{n} \left(\sigma_{ei}^2 + Var\left(\hat{\mu}_i\right)\right)$$

7.1.5 Proof of Proposition 4

To receive the $\overline{C}(\tau_m;\tau_n)$, I sum the equally weighted abnormal return at time t denoted $\overline{\varepsilon}_t^*$ over the period τ_m to τ_n. The length of this time interval is T_1.

$$\overline{C}(\tau_m;\tau_n) = \sum_{t=\tau_m}^{\tau_n} \overline{\varepsilon}_t^*$$

Assuming that the abnormal return of the portfolio is uncorrelated and the variances of daily portfolio weighted abnormal returns remain constant over time, I can derive the variance of $\overline{C}(\tau_m;\tau_n)$ in a simple manner.

$$Var\left(\overline{C}(\tau_m;\tau_n)\right) = \overline{\sigma}^2 = n^{-2} \cdot T_1 \cdot tr\left(\mathbf{I}_{n\times n}\boldsymbol{\sigma}_{\mathbf{e}}^{\mathbf{2}} + \mathbf{I}_{n\times n} Var\left(\hat{\boldsymbol{\mu}}\right)\right)$$

Thus, the test statistic can be written as…

$$\frac{\overline{C}(\tau_m;\tau_n)}{\overline{\sigma}(\tau_m;\tau_n)} = \frac{\sum_{t=\tau_m}^{\tau_n} \overline{\varepsilon}_t^*}{\sqrt{n^{-2} \cdot T_1 \cdot tr\left(\mathbf{I}_{n\times n}\boldsymbol{\sigma}_{\mathbf{e}}^{\mathbf{2}} + \mathbf{I}_{n\times n} Var\left(\hat{\boldsymbol{\mu}}\right)\right)}}$$

It is straightforward to show that this pivot variable is standard normally distributed, using the result provided by proposition 1.

7.2 Statistical appendix

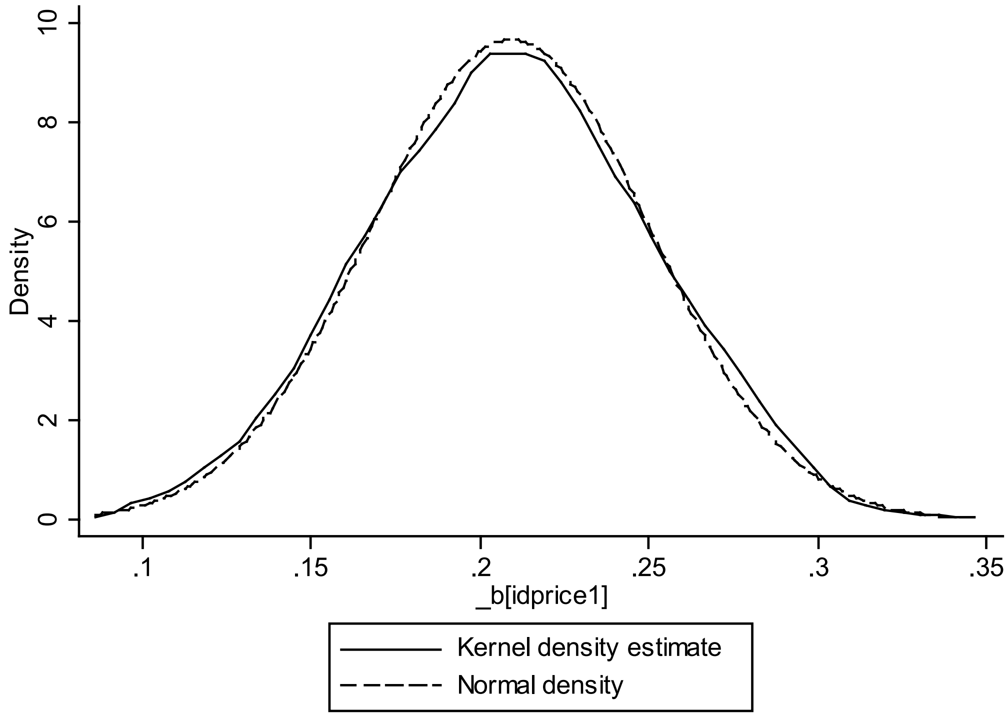
7.2.1 Bootstrapping to construct confidence intervals for impulse response functions

Because I stick to the method proposed by Lütkepohl (2000), a detailed technical description seems to be redundant. Nevertheless, interim results like bootstrapping distributions of impact multipliers are presented accompanied by some practical hints for STATA 8.0 users. Obviously, programming STATA means that a variety of different roads may lead to the same goal; hence, my proposed programs are just one way to think about solutions. Furthermore, I am aware of the fact that less complicated programs might exist.

First, the standard panel VAR with unexpected macroeconomic shocks (5.7) is estimated using OLS. Based on the residuals, share prices and dividends are reconstructed recursively, and the panel VAR model is re-estimated. This loop is conducted 1000 times and, thus, 1000 different point estimates of all coefficients are obtained. To illustrate this step, figure 7.1 depicts the distribution of estimated coefficients of Δp_{it-1}, the lagged first difference in share prices.

Figure 7.1: Kernel density of the coefficient of Δp_{it-1} and normal density

I use the Kernel density to illustrate the distribution of estimated coefficients; thereby, Epanechnikov function is used without any additional weights like frequency weights. To compare both distributions, figure 7.1 also plots the normal density.



All estimated coefficients are asymptotically normally distributed like the depicted example. The following step requires to calculate impact multipliers for different time horizons as shown in (5.9). Note that the matrix T contains the ordering restriction; hence, the impact multipliers reflect the Cholesky decomposition and the structural VAR model. After defining matrix T, the following STATA sequence yield the impact multipliers up to ten periods after an exogenous shock.

STATA program 7.1

```
*Facilitate bootstrapping procedure by using partitioned matrixes
set matsize 2000
mkmat idprice1, matrix(p1)
mkmat idprice2, matrix(p2)
mkmat idprice3, matrix(p3)
mkmat idprice4, matrix(p4)
mkmat idprice5, matrix(p5)
mkmat idprice6, matrix(p6)
mkmat iddiv1, matrix(d1)
mkmat iddiv2, matrix(d2)
mkmat iddiv3, matrix(d3)
mkmat iddiv4, matrix(d4)
mkmat iddiv5, matrix(d5)
mkmat iddiv6, matrix(d6)
mkmat v, matrix(v)
mkmat w, matrix(w)
mkmat b_idprice1, matrix(dp1)
mkmat b_idprice2, matrix(dp2)
mkmat b_idprice3, matrix(dp3)
mkmat b_idprice4, matrix(dp4)
mkmat b_idprice5, matrix(dp5)
mkmat b_idprice6, matrix(dp6)
mkmat b_iddiv1, matrix(dd1)
mkmat b_iddiv2, matrix(dd2)
mkmat b_iddiv3, matrix(dd3)
mkmat b_iddiv4, matrix(dd4)
mkmat b_iddiv5, matrix(dd5)
mkmat b_iddiv6, matrix(dd6)
mkmat b_v, matrix(dv)
mkmat b_w, matrix(dw)
```

Estimated coefficients stored as variables are transferred into column vectors

```
matrix delta=0.053881/0.394383
*Now define coefficient matrix A, B, and C, D
forvalues i=1/1000 {
     matrix define A`i'=[p1[`i',1], d1[`i',1] \ dp1[`i',1], dd1[`i',1]]
     matrix define B`i'=[p2[`i',1], d2[`i',1] \  dp2[`i',1], dd2[`i',1]]
     matrix define C`i'=[p3[`i',1], d3[`i',1] \ dp3[`i',1], dd3[`i',1]]
     matrix define D`i'=[p4[`i',1], d4[`i',1] \  dp4[`i',1], dd4[`i',1]]
     matrix define E`i'=[p5[`i',1], d5[`i',1] \  dp5[`i',1], dd5[`i',1]]
     matrix define F`i'=[p6[`i',1], d6[`i',1] \  dp6[`i',1], dd6[`i',1]]
     }
```

I define coefficient matrixes for every simulated replication

```
*Now define constraint matrix T
matrix T=[1, -delta[1,1] \ 0, 1]
*Derive impact multipliers
forvalues i=1/1000 {
     matrix I0=inv(T)
     matrix I1n`i'=A`i'*inv(T)
     matrix I2n`i'=A`i'*I1n`i'+B`i'*I0
     matrix I3n`i'=A`i'*I2n`i'+B`i'*I1n`i'+C`i'*I0
     matrix I4n`i'=A`i'*I3n`i'+B`i'*I2n`i'+C`i'*I1n`i'+D`i'*I0
     matrix I5n`i'=A`i'*I4n`i'+B`i'*I3n`i'+C`i'*I2n`i'+D`i'*I1n`i'+E`i'*I0
     matrix I6n`i'=A`i'*I5n`i'+B`i'*I4n`i'+C`i'*I3n`i'+D`i'*I2n`i'+E`i'*I1n`i'+F`i'*I0
     matrix I7n`i'=A`i'*I6n`i'+B`i'*I5n`i'+C`i'*I4n`i'+D`i'*I3n`i'+E`i'*I2n`i'+F`i'*I1n`i'
     matrix I8n`i'=A`i'*I7n`i'+B`i'*I6n`i'+C`i'*I5n`i'+D`i'*I4n`i'+E`i'*I3n`i'+F`i'*I2n`i'
     matrix I9n`i'=A`i'*I8n`i'+B`i'*I7n`i'+C`i'*I6n`i'+D`i'*I5n`i'+E`i'*I4n`i'+F`i'*I3n`i'
     matrix I10n`i'=A`i'*I9n`i'+B`i'*I8n`i'+C`i'*I7n`i'+D`i'*I6n`i'+E`i'*I5n`i'+F`i'*I4n`i'
     }
exit
```

Impact multipliers are determined using (5.9)

Based on this procedure, 1000 different impact multipliers for time horizon zero (immediate impact) to ten are determined. The impact of a sudden change in net national product can be expressed in the following manner.

STATA program 7.2

```
forvalues i=1/1000 {
    matrix define g`i'=[v[`i',1] \ dv[`i',1]]
    matrix J0n`i'=I0*g`i'
    forvalues j=1/10 {
        matrix J`j'n`i'=I`j'n`i'*g`i'
        }
    }
*Inflation has only a direct influence on prices
forvalues i=1/1000 {
    matrix define h`i'=[w[`i',1] \ 0]
    matrix H0n`i'=I0*h`i'
    forvalues j=1/10 {
        matrix H`j'n`i'=I`j'n`i'*h`i'
        }
    }
exit
```

The results in program 7.1 are used to derive the impact of a change in NNP

In a similar way, one can define the influence of unexpected changes in inflation rates

Applying program 7.2 yields 1000 matrixes (index `i’) H`j’`i´ for every time horizon one to ten (expressed by index `j’). Obviously, the column vector denoted H captures the influence of changes in inflation rates on share prices and dividends. The task is to isolate both dynamic responses and to store the resulting vectors as variables in order to derive descriptive statistics, namely the percentiles. Despite its mathematical simplicity, I have to use a very tricky program to capture the effect of inflation rates on share prices.

STATA program 7.3

```
*Now lets analyze the impact of inflation on prices
gen i=_n
tsset i
forvalues j=0/10 {
     forvalues i=1/1000 {
          matrix define i21n`i'`j'=H`j'n`i'[1,1]
          svmat i21n`i'`j', names(i21n`i'`j')
          rename i21n`i'`j'1 i21n`i'`j'
          }
     *Generate variable i21
     gen i21`j'=.
     *Trick to shift observations with lead structure
     forvalues i=1/1000 {
          qui gen i21l`i'`j'=i21n`i'`j'[_n-(`i'-1)]
          }
     forvalues i=1/1000 {
          qui replace i21`j' = i21l`i'`j' if i21`j'==. & i==`i'
          }
     forvalues i=1/1000 {
          qui matrix drop i21n`i'`j'
          qui drop i21n`i'`j'
          qui drop i21l`i'`j'
          }
     }
forvalues j=0/10 {
     sum i21`j', detail
     }
exit
```

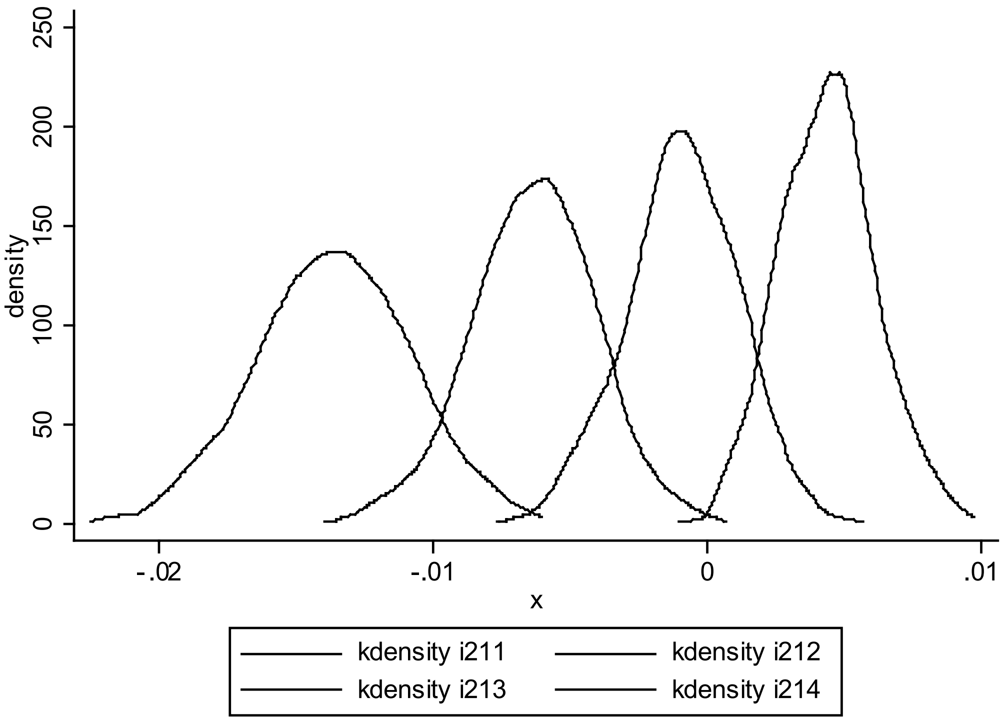
I define a matrix labeled i21 that contains the first element of matrix H

Note that I have defined 1000 variables labeled i21, but these variables have to be stored in one column vector. Every variable i21 has one observation if i (time index) is equal to one and 999 missing values. To combine these 1000 variables, I have to shift the single observations to avoid missing value problems

Consequently, it is possible to describe the distribution of these impact multipliers by applying kernel densities as shown in figure 7.2. The standard deviation becomes smaller with increasing time horizon; hence, the bootstrapping intervals become narrower. To construct bootstrapping intervals, the *sum* command in program 7.3 indicates the percentiles.

Figure 7.2: Kernel density of the 1000 simulated impact multipliers

I use the Kernel density to illustrate the distribution of impact multipliers; thereby, the dynamic response of share prices due to inflation shocks is measured. To illustrate the response over time, figure 7.2 plots the distributions one to four periods after the exogenous shock.



7.2.2 Decomposing time series into transitory and permanent components

Deriving coefficient matrix A_i for an individual firm respectively an industry is straightforward using (5.22) and simple matrix calculus; hence I skip this part of the program. Applying the *eigenvalues* STATA command yields the eigenvalues of matrix A_i; thereby, the eigenvalues are sorted as highlighted in (5.25). Afterwards, I build up the non-linear system of equations (5.26). Program 7.4 shows the details of this step.

STATA program 7.4

```
matrix eigenvalues r c = A ;

matrix list r ;

matvsort r l, decrease;

matrix list l;

matrix M1=A-l[1,1]*I(3);

matrix M2=A-l[1,2]*I(3);

matrix M3=A-l[1,3]*I(3);
```

M1 to M2 are the first terms in (5.26) and are the difference between matrix A_i (5.22) and the respective eignevalue times identity matrix denoted I(3). Setting 3 in parentheses defines that identity matrix I is 3×3 dimensional.

```
#delimit cr

program define nlfaq1
  if "`1'"=="?" {
     global S_1 "A B C"
     global A=.1
     global B=.1
     global C=.1
     exit
  }
     tempvar yh
     gen `yh'=M1[1,1]*$A+M1[1,2]*$B+M1[1,3]*$C+1 in 1
     replace `yh'= M1[2,1]*$A+M1[2,2]*$B+M1[2,3]*$C in 2
     replace `yh'=M1[3,1]*$A+M1[3,2]*$B+M1[3,3]*$C in 3
     replace `yh'=$A*$A+$B*$B+$C*$C-1 in 4
     replace `1'=`yh'
end
```

System of equations (5.26) in STATA syntax

```
input y
     1
     0
     0
     0
     end
```

I define an artifical dependent variable denoted y.

```
nl faq1 y
exit
```

Using numerical methods yield the solution for the first eigenvector k_{i1}.

By slightly modifying program 7.4, the other two eigenvectors k_{i2} and k_{i3} can be calculated. The last steps are straightforward and, thus, are not discussed precisely. After determining the eigenvectors, one has to combine the results in the partitioned matrix K_i as described in (5.23). Then, canonical variates η_{it} are obtained following definition (5.24) and tested by applying unit-root tests.

8. References

8.1 Cited literature

Abuaf, N. and P. Jorion, 1990, Purchasing power parity in the long run, Journal of Finance 45, 157-174.

Anari, A. and J. Kolari, 2001, Stock prices and inflation, Journal of Financial Research 24, 587-602.

Andrews, D. W. K., 1993, Tests for parameter instability and structural change with an unknown change point, Econometrica 61, 821-856.

Andrews, D. W. K. and E. Zivot, 1992, Further evidence on Great Crash, the oil price shock and the unit root hypothesis, Journal of Business and Economic Statistics 10, 251-270.

Arellano, M. and O. Bover, 1995, Another look at the instrumental variable estimation of error-component models, Journal of Econometrics 10, 29-51.

Armitage, S., 1995, Event study methods an evidence on their performance, Journal Economic Surveys 9, 25-52.

Asquith, P., 1983, Merger bids, uncertainty, and stockholder returns, Journal of Financial Economics 11, 51-83.

Bachmann, H., 1898, Die Effektenspekulation mit besonderer Berücksichtigung der Ergebnisse der Deutschen Börsenenquete, Zeitschrift für Schweizerische Statistik 34, 248-288.

Ball, R. and Brown, P., 1968, An empirical evaluation of accounting income numbers." Journal of Accounting Research 6, 159-178.

Baltzer, M. and G. Kling, 2003, Resiliency of the pre-World War I German stock exchange: Evidence from a panel vector autoregression, conference proceedings of the 5th World Congress of Cliometrics 2004, 169-180.

Banerjee, A. and E.W. Eckard, 2001, Why regulate insider trading? Evidence from the first Great Merger Wave (1896-1903), The American Economic Review 91, 1329-1349.

Banerjee, A., Lumsdaine, R. L. and J. H. Stock, 1992, Recursive and sequential tests of the unit-root and trend-breaking hypotheses: Theory and international evidence, Journal of Business & Economic Statistics 10, 271-287.

Barclay, M. and R. Litzenberger, 1988, Announcement effects of new equity issues and the use of intraday price data, Journal of Financial Economics 21, 71-100.

Baten, J., 2001a, Produktivitätsvorteil in kleinen und mittelgroßen Industrieunternehmen, Sicherheit in Großunternehmen? Die Gesamtfaktorproduktivität um 1900, Tübinger Diskussionsbeitrag 217.

Baten, J., 2001b, Große und kleine Unternehmen in der Krise von 1900-1902, Tübinger Diskussionsbeitrag 216.

Baten, J., 2001c, Expansion und Überleben von Unternehmen in der `Ersten Phase der Globalisierung´, Tübinger Diskussionsbeitrag 215.

Beck, N. and J. N. Katz, 1995, What to do (and not to do) with time series cross-sectional data, American Policy Science Review 89, 634-647.

Berghoff, H., 2002, Der Berliner Kapitalmarkt im Aufbruch (1830-1870), in: Geschichte des Finanzplatzes Berlin, Pohl, H. (ed.), Fritz Knapp Verlag, Frankfurt am Main, 53-102.

Berle, A. A. and G. C. Means, 1932, The modern corporation and private property, Commerce Clearing House, Chicago.

Bernard, V.L., 1987, Cross-sectional dependence and problems in inference in market-based accounting research, Journal of Accounting Research 25, 1-48.

Binder, J., 1985, On the use of the multivariate regression model in event studies, Journal of Accounting Research 23, 370-383.

Black, F., 1986, Noise, Journal of Finance 41, 529-543.

Blaich, F., 1973, Kartell- und Monopolpolitik im kaiserlichen Deutschland. Das Problem der Marktmacht im deutschen Reichstag zwischen 1879 und 1914, Düsseldorf.

Blanchard, O. J. and D. Quah, 1989, The dynamic effects of aggregate supply and demand disturbances, American Economic Review 79, 655-673.

Böckem, S., 2002, Merger mania or after you - a theory of endogenous merger waves, Working Paper, Universität Dortmund.

Boehmer, E. ; Musumeci, J. and Poulsen, A., 1991, Event- study methodology under conditions of event-induced variance, Journal of Financial Economics 30, 253-272.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics 31, 307-327.

Bond, S. and C. Meghir, 1994, Dynamic investment models and firm’s financial policy, The Review of Economic Studies 61, 197-222.

Borchardt, K. and A. Rischl, 1991, Inflation and relative price variability under a gold standard: Evidence from Germany (1850-1913): A Comment, Jahrbücher für Nationalöknomie und Statistik 208, 541-545.

Bordo, M, Eichengreen, B, and D. Irwin, 1999, Is globalization today really different than globalization a hundred years ago?, in: Brookings trade forum 1999, Collins, S. M. and R. Z. Lawrence (ed.), Brookings Institute Press, Washington.

Borg, J.R. and Leeth, J.D., 2000, The impact of takeovers on shareholder wealth during the 1920s merger wave, Journal of Financial and Quantitative Analysis 35, 217-238.

Borg, J.R. and Leeth, J.D., 1994, The impact of mergers on acquiring firm shareholder wealth: The 1905-1930 experience, Empirica 21, 221-244.

Borg, J. R. ; Borg, M. O. and Leeth, J.D., 1989, The success of mergers in the 1920s: A stock market appraisal of the second merger wave, International Journal of Industrial Organization 7, 117-131.

Brown, C. and J. L. Medoff, 1987, The impact of firm acquisitions on labor, NBER working paper: 2273.

Brown, S. J. and Warner, J.B., 1985, Using daily stock returns: The case of event studies, Journal of Financial Economics 14, 3-31.

Brown, S. J. and Warner, J.B., 1980, Measuring security price performance, Journal of Financial Economics 8, 205-258.

Brown, R. L., Durbin, J. and J. M. Evans, 1975, Techniques for testing the constancy of regression relationships over time, Journal of the Royal Statistical Society, Series B, 37, 149-192.

Brudney, V., 1979, Insiders, outsiders, and informational advantages under the Federal Securities Law, Harvard Law Review 93, 322-376.

Bühner, R., 1991, The success of mergers in Germany, International Journal of Industrial Organization 9, 513-553.

Burhop, C. and G. Wolff, 2003, A compromise estimate of the net national product and the business cycle in Germany, 1851 – 1913, conference proceedings of the 5th World Congress of Cliometrics 2004, 97-108.

Burhop, C. and G. Wolff, 2002, National accounting and the business cycle in Germany, 1851-1913, unpublished working paper, http://www.uni-muenster.de/Geschichte Philosophie/ Geschichte/ hist-sem/ SW-G/ forschung/ papers/ kaiser.pdf.

Campbell, J. Y. and R. J. Shiller, 1988, Stock prices earnings, and expected dividends, Journal of Finance 43, 661-676.

Cheung, Y. W. and K. S. Lai, 1993, Finite-sample size of Johansen's likelihood ratio tests for cointegration, Oxford Bulletin of Economics and Statistics 55, 313-328.

Chan, L. and J. Lakonishok, 1995, The behavior of stock prices around institutional trades, Journal of Finance 50, 1147-1174.

Chandler, A. D., 1990, Scale and scope: The dynamics of industrial capitalism, Cambridge.

Chandler, A. D., 1977, The visible hand, Cambridge.

Cermeño, R. and K. B. Grier, 2001, Modeling GARCH processes in panel data: Theory, simulations and examples, unpublished working paper, http://www.ou.edu/cas/econ/research/2001-3.pdf

Collins, D. W. and W. T. Dent, 1984, A comparison of alternative testing methodologies used in capital market research, Journal of Accounting Research 22, 48-84.

Corrado, C. J., 1989, A nonparametric test for abnormal security-price performance in event studies, Journal of Financial Economics 23, 385-395.

Corrado, C. J. and Zivney, T.L., 1992, The specification and power of the sign test in event study hypothesis tests using daily stock returns, Journal of Financial and Quantitative Analysis 27, 465-478.

Cosh, A. D. and P. M. Guest, 2001, The long-run performance of hostile takeovers: U.K. evidence, presented at the *EARIE* conference in Dublin.

Davis, L., Neal, L, and E. White, 2002, The effect of deflation on the first global capital market: The financial crises of the 1890s and the responses of the stock exchanges in London, New York, Paris, and Berlin, unpublished working paper, http://www.eh.net/XIIICongress/cd/papers/45Neal400.pdf.

Davis, L., 1966, The capital markets and industrial concentration; the US and the UK, a comparative study, Economic History Review 19, 255-272.

Day, R. A., 1999, How to write & publish a scientific paper, 5th edition, Oryx Press, Westport.

Day, R. A., 1995, Scientific English: A guide for scientists and other professionals, 2nd edition, Oryx Press, Westport.

De Long, J. B. and M. Becht, 1992, “Excess Volatility” and the German stock market, 1876-1990, NBER Working Papers: 4054.

Denis, D. K. and J. J. McConnell, 1986, Corporate mergers and security returns, Journal of Financial Economics 16, 143-187.

Dibooglu, S., 1998, The incidence and effects of macroeconomic disturbances under alternative exchange rate systems: Evidence since the classical gold standard, International Review of Economics and Finance 7, 225-241.

Dimson, E. and P. Marsh, 1984, An analysis of brokers’ and analysts’ unpublished forecasts of UK stock returns, Journal of Finance 39, 1257-1292.

Dong, M.; Hirshleifer, D.; Richardson, S. and S. H. Teoh, 2003, Does investor misvaluation drive the takeover market?, Ohio State University, working paper, http://www.stern.nyu.edu/clb/hirshleifer.pdf.

Dyckman, T. ; Philbrick, D. and Stephan, J., 1984, A comparison of event study methodologies using daily stock returns: A simulation approach, Journal of Accounting Research 22, 1-33.

Easley, D. and M. O'Hara, 1992, Adverse selection and large trade volume: The implication for market efficiency, Journal of Financial and Quantitative Analysis 27, 185-208.

Eckbo, B., 1986, Mergers and the market for corporate control, Canadian Journal of Economics 19, 236-260.

Eis, C., 1971, The 1919-1930 merger movement in American industry, Arno Press, New York.

Elbaum, B. and W. Lazonick, 1986, An institutional perspective on the British decline, in: The decline of the British Economy, Elbaum, B. and W. Lazonick (ed.), Oxford.

Elfron, B. and R. J. Tibshirani, 1993, An introduction to the bootstrap, New York, Chapman & Hall.

Engle, R. F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987-1007.

Eube, S., 1998, Der Aktienmarkt in Deutschland vor dem Ersten Weltkrieg. Eine Indexanalyse, Frankfurt a. M.

Falke, J., 1979, Anlegerschutz und Unternehmenspublizität. Die Zulassung von Aktien zum Börsenhandel bis zur New Deal-Gesetzgebung in vergleichender Sicht, in: Recht und Entwicklung der Großunternehmen im 19. und frühen 20. Jahrhundert, Horn, N. and J. Kocka, Vandenhoeck & Ruprecht, Göttingen.

Fama, E. and French, K., 1988, Permanent and temporary components of stock prices, Journal of Political Economy 96, 246-273.

Fama, E., 1970, Efficient capital markets: A review of theory and empirical work, Journal of Finance 25, 383-417.

Fama, E. ; Fisher, L. ; Jensen, M. and Roll, R., 1969, The adjustment of stock prices to new information, International Economic Review 10, 1-21.

Farrell, J. and C. Shapiro, 1990, Horizontal mergers: An equilibrium analysis, American Economic Review 80, 107-126.

Feldenkirchen, W., 1988, Concentration in German industry, 1870-1939, in: The concentration process in the entrepreneurial economy since the late 19th century, Pohl, H. (ed.), Steiner-Verlag, Wiesbaden.

Firth, M., 1980, Takeovers, shareholder returns, and the theory of the firm, Quarterly Journal of Economics 94, 235-260.

Fohlin, C., 2002, Regulation, taxation and the development of the German universal banking system, 1884-1913, European Review of Economic History 6, 221-254.

Fremdling, R., 1991, Productivity comparison between Great Britain and Germany, 1855-1913, Scandinavian Economic History Review 39, 28-42.

Fremdling, R. and J. Krengel, 1985, Kartelle und ihre volks- bzw. Einzelwirtschaftliche Bedeutung bis 1914, in: Kartelle und Kartellgesetzgebung in Praxis und Rechtsprechung vom 19. Jahrhundert bis zur Gegenwart, Pohl, H. (ed.), Stuttgart.

Gerhausser, K., 1990, Inflation and relative price variability under a gold standard: Evidence from Germany (1850-1913), Jahrbücher für Nationalöknomie und Statistik 207, 391-394.

Gerschenkron, A., 1962, Economic backwardness in historical perspective, Cambridge.

Gömmel, R., 1985, Kartelle in the öffentlichen Meinung bis 1914, in: Kartelle und Kartellgesetzgebung in Praxis und Rechtsprechung vom 19. Jahrhundert bis zur Gegenwart, Pohl, H. (ed.), Stuttgart.

Gregory, A. W. and B. E. Hansen, 1996, Residual-based tests for cointegration in models with regime shifts, Journal of Econometrics 70, 99-126.

Grossman, S. J. and J. E. Stiglitz, 1980, On the impossibility of informationally efficient markets, American Economic Review 70, 393-408.

Hadri, K., 2000, Testing for stationarity in heterogeneous panel data, The Econometrics Journal 3, 148-161.

Hamilton, J. D, 1994, Time series analysis, Princeton, Princeton University Press.

Han, A. K. and D. Park, 1989, Testing for structural change in panel data: Application to a study of U.S. foreign trade in manufacturing goods, The Review of Economics and Statistics 71, 135-142.

Hannah, L., 1974, Mergers in British manufacturing industry, 1880-1918, Oxford Economic Papers 26, 1-20.

Harris, R. and R. Sollis, 2003, Applied time series modeling and forecasting, John Wiley & Sons Ltd, Chichester, UK.

Hatton, T. and J. Williamson, 1998, The age of mass migration: Causes and economic impact, Oxford University Press, New York.

Hatton, T. and J. Williamson, 1994, What drove the mass migration from Europe in the late nineteenth?, Population and Development Review 20, 533-559.

Henning, F. W., 1973, Die Industralisierung in Deutschland 1800 bis 1914, Paderborn.

Ho, T., 2002, A panel cointegration approach to the investment-saving correlation, Empirical Economics 27, 91-100.

Hoffmann, W. G. et al., 1965, Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts, Berlin.

Holden, C., and A. Subrahmanyam, 1992, Long-lived private information and imperfect competition, Journal of Finance 47, 247-270.

Hsiao, C., 1992, The econometrics of panel data: Handbook of theory and applications, Advanced Studies in Theoretical and Applied Econometrics, vol. 28. Norwell, Mass. and Dordrecht: Kluwer Academic.

Huerkamp C., 1979, Fusionen der 100 größten Unternehmen von 1907 zwischen 1887 und 1907, in: Horn and Kocka (ed.), Recht und Entwicklung der Grossunternehmen, Göttingen.

Im, K., Pesaran, M. H. and Y. Shin, 1997, Testing for unit roots in heterogeneous panels. unpublished working paper, Department of Applied Economics, University of Cambridge.

Jacobs, A. and H. Richter, 1935, Die Großhandelspreise in Deutschland von 1792 bis 1934, Hamburg.

Jarrell, G. A. and A. B. Poulsen, 1989, The returns to acquiring firms in tender offers: Evidence from three decades, Financial Management 18, 12-19.

Jarrell, G. A. ; Brickley, J. A. and J. M. Netter, 1988, The market for corporate control: The empirical evidence since 1980, Journal of Economic Perspective 2, 49-68.

Jennings, R., 1994, Intraday changes in target firms` share price and bid-ask quotes around takeover announcements, Journal of Financial Research 17, 255-270.

Jensen, M. C., 1986, Agency costs of free cash flow, corporate finance, and takeovers, American Economic Review 76, 323-329.

Jensen, M. C. and Ruback, R.S., 1983, The market for corporate control: The scientific evidence, Journal of Financial Economics 11, 5-50.

Jochum, C., G. Kirchgässner and M. Platek, 1999, A long-run relationship between Eastern European stock markets? Cointegration and the 1997/98 crisis in emerging markets, Weltwirtschaftliches Archiv 135, 454-479.

Kennedy, P., 1998, A guide to econometrics, fourth edition, Blackwell Publishers, Oxford.

Keown, A. J. and Pinkerton, J. M., 1981, Merger announcements and insider trading activity: An empirical investigation, Journal of Finance 36, 855-869.

Kim, O. and Verrecchia, R.E., 1994, Market liquidity and volume around earnings announcements, Journal of Accounting and Economics 17, 41-67.

Kinghorn, J. R. and J. V. Nye, 1996, The scale of production in western economic development: A comparison of official industry statistics in the United States, Britain, France, and Germany, 1905-1913, Journal of Economic History 56, 90-112.

Kleidon, A. W., 1986, Variance bound tests and stock price valuation models, Journal of Political Economy 94, 953-1001.

Klein, A. and J. Rosenfeld, 1987, The influence of market conditions on event-study residuals, Journal of Financial and Quantitative Analysis 22, 345-351.

Kling, G., 2003a, Disclosure of mergers without regulatory restrictions: Comparing insider-trading in the year 1908 and 2000 in Germany, conference proceedings of the annual conference of the Economic History Society 2004, 60-64.

Kling, G., 2003b, The impact of merger announcements on stock prices: The rejection of the merger paradox for German companies, presented at the *Economics and Business Historical Society Conference* 2003 in Memphis and at the *European Historical Economics Society Conference* 2003 in Madrid.

Kling, G., 2002a, The impact of trading mechanisms on liquidity, order processing costs, and the degree of information asymmetry, unpublished working paper.

Kling, G., 2002b, Information asymmetry around earnings announcements: A simple model to decompose the bid-ask spread, unpublished working paper.

Kocka, J. and H. Siegrist, 1979, Die hundert größten Industrieunternehmen im späten 19. und frühen 20.Jahrhundert. Expansion, Diversifikation und Integration im internationalen Vegleich, in: Recht und Entwicklung der Großunternehmen im 19. und frühen 20.Jahrhundert, Horn, N. and J. Kocka (ed.), Vandenhoeck&Ruprecht, Göttingen.

Kloy, J. W., Schindler, K., and J. Welcker, 1992, Professionelles Optionsgeschäft, third edition, Zürich.

Knight, F. H., 1921, Risk, uncertainty and profit, Houghton Mifflin, Houston.

Kramer, W. and W. Ploberger, 1992, The CUSUM test with OLS residuals, Econometrica 60, 271-285.

Krengel, J., 1982, Die deutsche Roheisenindustrie 1871-1913. Eine quantitative-historische Untersuchung, Dissertation, FU Berlin.

Kyle, A. S., 1985, Continuous auctions and insider trading, Econometrica 53, 1315-1335.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?, Journal of Econometrics 54, 159-178.

Larsson, R., Lyhagen, J. and M. Lothgren, 2001, Likelihood-based cointegration tests in heterogeneous panels, Econometrics Journal 4, 109-142.

Lee, B., 1998, Permanent, temporary, and non-fundamental components of stock prices, Journal of Financial and Quantitative Analysis 33, 1-32.

Leland, H. E., 1992, Insider trading: Should it be prohibited?, Journal of Political Economy 100, 859-887.

Leland, H. and D. Pyle, 1977, Information asymmetries, financial structure and financial intermediaries, Journal of Finance 32, 371-387.

Levin, A., Lin, C. and C. J. Chu, 2002, Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties, Journal of Econometrics 108, 1-24.

Levin, D., 1990, Horizontal mergers: The 50-percent benchmark, American Economic Review 80, 1238-1245.

Levin, I. P., 1999, Relating statistics and experimental design, Sage Publications, Thousand Oaks.

Li, L. and Z. Hu, 1998, Responses of the stock market to macroeconomic announcements across economic states, International Monetary Fund, working paper.

Liu, Y. A. and M. Pan, 2003, Stock price volatility and permanent components of earnings and dividends,unpublished working paper.

Lo, A. and A. C. MacKinlay, 1990, An econometric analysis of nonsynchronous-trading, Journal of Econometrics 45, 181-212.

Love, I. and L. Zicchino, 2003, Financial development and dynamic investment behavior: Evidence from panel vector autoregression, unpublished working paper, http://econ.worldbank.org/files/20826_wps2913.pdf.

Lütkepohl, H., 2000, Bootstrapping impulse responses in VAR Analyses, Berlin.

Madsen, J. B., 2002, The share market boom and the recent disinflation in the OECD countries: The tax-effects, the inflation-illusion and the risk-aversion hypotheses reconsidered, Quarterly Review of Economic and Finance 42, 115-141.

Malatesta, P., 1986, Measuring abnormal performance: The event parameter approach using joint generalized least squares, Journal of Financial and Quantitative Analysis 21, 27-38.

Mandelker, G., 1974, Risk and return: The case of merging firms, Journal of Financial Economics 1, 303-335.

Magenheim, E. B. and Mueller, D. C., 1988, On measuring the effect of acquisitions on acquiring firm shareholders', in Coffee, J.C., Lowenstein, L. and S. Rose-Ackerman (eds.), Knights, raiders, and targets: The impact of the hostile takeover, New York, NY, Oxford University Press.

Manne, H. G., 1965, Mergers and the market for corporate control, Journal of Political Economy 73, 110-120.

Masulis, R. W., 1980, The effect of capital structure change on security prices, Journal of Financial Economics 8, 139-177.

McCloskey, D. N., 1994, Why don't economists believe empirical findings?, Eastern Economic Journal 20, 357-360.

McFadden, D., 1974, The measurement of urban travel demand, In Zarembka, P. (ed.), Frontiers in econometrics, New York, Academic Press.

Mac Kinnon, J. G., 1994, Approximate asymptotic distribution functions for unit-root and cointegration tests, Journal of Business and Economic Statistics 12, 167-176.

Meltzer, A. H. and S. Robinson, 1988, Stability under the Gold Standard in Practice, in M. D. Bordo (ed.), Money, History, and International Finance: Essays in Honor of Anna J. Schwartz, Chicago, 163-195.

Mills, T. C., 1990, Time series techniques for economists, Cambridge University Press.

Modigliani, F. and R.A. Cohn, 1979, Inflation, rational valuation and the market, Financial Analysts Journal 35, 24-44.

Morse, D. , 1984, An econometric analysis of the choice of daily versus monthly returns in tests of information content, Journal of Accounting Research 22, 605-623.

Myers, S. C. and N. J. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, Journal of Financial Economics 13, 187-221.

Najand, M. and H. Rahman, 1991, Stock market volatility and macroeconomic variables: International evidence, Journal of Multinational Financial Management 1, 51-66.

Nelson, R. L., 1959, Merger Movements in American Industry 1895-1956, Princeton University Press; Princeton, New Jersey.

Newey, W.K. and K.D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703-708.

Palacios-Huerta, I., 2003, Professionals play minimax, Review of Economic Studies 70, 395-415.

Pedroni, P., 2001, Purchasing power parity tests in cointegrated panels, Review of Economics and Statistics 83, 727-731.

Pedroni, P., 2000, Fully modified OLS for heterogeneous cointegrated panels, in Baltagi, B. H., Fomby, T. B., and R. C. Hill (eds), Nonstationary panels, panel cointegration, and dynamic panels, Advances in Econometrics, Vol.15, Elsevier Science, Amsterdam.

Pedroni, P., 1999, Critical values for cointegration tests in heterogeneous panels with multiple regressors, Oxford Bulletin of Economics and Statistics 61, 653-670.

Perron, P., 1994, Trend, unit root and structural change in macroeconomic time series, in B. B. Rao (ed.), Cointegration for applied economists, Macmillian Press, Basingtoke, UK.

Perron, P., 1989, The Great Crash, the oil shock and the unit root hypothesis, Econometrica 57, 1361-1402.

Pitelis, C. N., 2004, Corporate governance, shareholder value and sustainable economic performance, Corporate Governance 12, 210-223.

Pound, J. and R. J. Zeckhauser, 1990, Clearly heard on the street: The effect of takeover rumors on stock prices, Journal of Business 63, 291-308.

Rapach, D. E., 2002, The long-run relationship between inflation and real stock prices, Journal of Macroeconomics 24, 331-351.

Rapach, D. E., 2001, Macro shocks and real stock prices, Journal of Economics and Business 53, 5-26.

Reimers, H. E., 1991, Analyse kointegrierter Variablen mittels autoregressiver Modelle, Heidelberg, Physica-Verlag.

Ritter, J. and Warr, R. S., 2002, The decline of inflation and the bull market of 1982-1999, Journal of Financial and Quantitative Analysis 37, 29-61.

Roll, R., 1986, Empirical evidence on takeover activity and shareholder wealth, in: Coffee, J.; Lowenstein, L. and S. Rose-Ackerman (ed.), Takeovers and contests for corporate control, Oxford University Press, New York.

Roll, R., 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, Journal of Finance 39, 1127-1139.

Rydqvist, K., 1996, Takeover bids and the relative prices of shares that differ in their voting rights, Journal of Banking and Finance 20, 1407-1425.

Sachs, J. and A. Warner, 1996, Economic reform and the process of global integration, Development Discussion Paper 552, Harvard Institute of International Development.

Salant, S., Switzer, S. and R. Reynolds, 1983, Losses from horizontal mergers: The effect of an exogenous change in industry structure on Cournot-Nash equilibrium, Quarterly Journal of Economics 98, 185-199.

Sarno, L. and M. P. Taylor, 1998a, Real exchange rates under the current float: unequivocal evidence of mean reversion, Economics Letters 60, 131-137.

Sarno, L. and M. P. Taylor, 1998b, The behavior of real exchange rates during the post-Bretton Woods period. Journal of International Economics 46, 281-312.

Savickas, R., 2003, Event-induced uncertainty and tests for abnormal performance, Journal of Financial Research 26, 165-178.

Schenk, H., 2001, Mergers and the economy: Theory and policy implications, presented at the workshop entitled `European Integration, Financial Systems and Corporate Performance´, UN University, Institute for New Technologies Kasteel Vaeshartelt.

Scherer, F. M., 1988, Corporate takeovers: The efficiency arguments, Journal of Economic Perspective 2, 69-82.

Sefcik, S. and Thompson, R., 1986, An approach to statistical inference in cross-sectional models with security abnormal returns as dependent variable, Journal of Accounting Research 24, 316-334.

Shiller, R., 1981, Do stock prices move too much to be justified by subsequent changes in dividends?, American Economic Review 71, 421-426.

Shleifer, A. and R. W. Vishny, 1988, Value maximization and the acquisition process, Journal of Economic Perspective 2, 7-20

Shleifer, A. and L. Summers, 1987, Hostile takeovers as breaches of trust, NBER working paper.

Smith, A., 1776, An enquiry into the nature and causes of the wealth of nations, in Campbell, R. H. and A. S. Skinner (eds.), 1976, Clarendon Press, Oxford.

Stigler, G. J., 1950, Monopoly and oligopoly by merger, The American Economic Review 40, 23-34.

Tatom, J. A., 2002, Stock prices, inflation and monetary policy, Business Economics 37, 7-19.

Tilly, R., 1999, Globalisierung aus historischer Sicht und das Lernen aus der Geschichte, Köln.

Tilly, R., 1992, Der deutsche Kapitalmarkt und die Auslandsinvestitionen von 1870 bis 1913, in: ifo Studien 38, 200-225.

Tilly, R., 1982, Mergers, external growth, and finance in the development of large-scale enterprise in Germany, 1880-1913, Journal of Economic History 42, 629-658.

Tilly, R., 1986, German banking, 1850-1914. Development assistance for the strong, Journal of European Economic History 15, 113-52.

Tilly, R., 1980, Banken und Industrialisierung in Deutschland: Quantifizierungsversuche, in: Entwicklung und Aufgaben von Versicherungen und Banken in der Industrialisierung, Henning, F. W.(ed.), Berlin.

Tsay, R. S. and G. C. Tiao, 1990, Asymptotic properties of multivariate Nonstationary processes with application to autoregressions, Annals of Statistics 18, 220-250.

Travlos, N. G. and Papaioannou, G.J., 1991, Corporate acquisitions: Method of payment effects, capital structure effects, and bidding firms` stock returns, Quarterly Journal of Business and Economics 30, 3-22.

Ukaegbu, E. O., 1987, Motives for corporate mergers and takeovers: An investigation of the `failing company´ hypothesis and of post-merger performance, PhD thesis, University of Stirling.

Strunk, W. and E. B. White, 1999, The elements of style, 4th edition, Allyn & Bacon, Needham Heights.

Van Huyck, J. B., 1987, A retrospective on the classical gold standard, 1821-1931: A review essay, Journal of Monetary Economics 19, 451-456.

Vásquez, I., 2000, The return to a global economy, Ideas on Liberty.

Warshow, H. T., 1924, The distribution of corporate ownership in the United States, Quarterly Journal of Economics 39, 15-38.

Weber, M., 2000, Börsenwesen: Schriften und Reden 1893-1898, in Borchardt, K., ed.: Max Weber: Börsenwesen (J.C.B. Mohr (Paul Siebeck) Tuebingen).

Wengenroth, U, 1985, Die Entwicklung der Kartellbewegung bis 1914, in: Kartelle und Kartellgesetzgebung in Praxis und Rechtsprechung vom 19. Jahrhundert bis zur Gegenwart, Pohl, H. (ed.), Stuttgart.

West, K. D., 1988, Bubbles, fads and stock price volatility tests: A partial evaluation, Journal of Finance 43, 639-656.

Wetzel, C., 1996, Die Auswirkungen des Reichsbörsengesetzes von 1896 auf die Effektenbörsen im Deutschen Reich, insbesondere auf die Berliner Fondsbörse (University of Muenster, dissertation).

Weston, J. F., 1953, The role of mergers in the growth of large firms, Berkeley.

Williamson, J. G., 1996, Globalization, convergence and history, Journal of Economic History 56, 277-306.

Winters, P. R., 1960, Forecasting Sales by Exponentially Weighted Moving Averages, Management Science 6, 324-342.

Wooldridge, J. M., 2002, Econometric analysis of cross section and panel data, MIT Press, Cambridge.

Zeileis, A., 2004, Alternative boundaries for CUSUM tests, Statistical Papers, forthcoming.

Zivney, T. L.; Bertin, W. J. and K. M. Torabzadeh, 1996, Overreaction to takeover speculation, Quarterly Review of Economics and Finance 36, 89-115.

8.2 Data sources and law texts

8.2.1 Law text

Börsengesetz vom 22.Juni 1896, Reichs-Gesetzblatt 1896, Number 15.

Börsengesetz vom 22.Juni 1896, in Pohl, H., ed.: Deutsche Börsengeschichte, 1992, 377-398.

8.2.2 Data sources

Berliner Börsenzeitung, issues for the years 1907 and 1908, available as microfilm.

Deutsche Bundesbank, 1976, Deutsches Geld- und Bankwesen in Zahlen 1876-1975, Frankfurt am Main.

Grabas, M., 1992, Konjunktur und Wachstum in Deutschland von 1895 bis 1914, Duncker & Humblot, Berlin.

Handbuch der deutschen Aktien-Gesellschaften, issue 1913-1914, volume I and II.

Kaiserliches Statistisches Amt, 1911/2, Vierteljahreshefte zur Statistik des Deutschen Reichs – Ergänzungsheft zu 1913, II, Berlin.

Saling's Börsen-Jahrbuch, 1913/4, Ein Handbuch für Kapitalisten und Effektenbesitzer. 2nd ed., Berlin, Leipzig.

Saling's Börsen-Papiere, 1874-6, vol. 1-5, 4th ed., Berlin.