Fama/French Factors for Germany: Which Set Is Best?

Roman Brückner, Patrick Lehmann, Martin H. Schmidt, Richard Stehle*

Abstract

We are aware of five internet-based data sets that provide Fama/French factors for Germany, in addition to the sets that we have created. They indicate that in the last twenty years a reverse size effect, a strong value effect and a very strong momentum effect existed.

We find that the provided time series for the market portfolio and the SMB, HML and MOM factors differ considerably, especially in the last century. We conclude that the major cause for the differences is the quality of the underlying stock market data bases. We also find that their quality seems to improve over time. During the last ten years the best factor data sets are very similar.

We first discuss the construction details, then compare the factors using standard procedures. In a next step, we use the factors to do a performance analysis of mutual funds, and finally to evaluate double-sorted portfolios. We find that the factor sets produce very different results, particularly for the mutual funds.

Our main advice for the users of Fama/French factor time series covering Germany and other countries is to check first, which factor data set fits best to their other data, then apply at least two factor sets. To the suppliers, we recommend to include a local expert for each country that is covered and to begin their work by carefully checking the quality of the underlying databases.

Version: 15 January 2014

Working Paper

---

* School of Business and Economics at Humboldt University Berlin, Dorotheenstr. 1, 10117 Berlin, Germany. Contact: stehle@wiwi.hu-berlin.de. We are grateful for valuable comments received from Ioana Sima and seminar participants at ESMT. We also acknowledge support by the Deutsche Forschungsgemeinschaft through the CRC 649 “Economic Risk”.

1 Introduction and Summary

The factors proposed by Fama/French (1993) and Carhart (1997) play an important role in the analysis of financial data, first in the U.S. and recently also "around the world". Important areas in which they are used include event studies, mutual fund performance studies, cost of equity capital estimations, and CAPM tests. Following Liew/Vassalou (2000) and Griffin (2002), Fama/French factors were computed for a selection of countries and used in a number of studies. For the U.S. capital market the calculation procedure suggested by Fama/French (1993) and the factor data supplied by Ken French are well accepted and used in most studies.[1] For other capital markets the situation is less clear. For some countries, to our knowledge, only one locally produced data set is freely available on the internet,[2] for other countries several.[3] In addition, several academics provide data sets for a number of countries (see below). In the UK, some alternative ways of calculating the SMB and HML factors were proposed and compared in recent papers, especially in Michou/Mouselli/Stark (2010) and Gregory/Tharyan/Christidis (2013). A similar discussion relating to the German market has not taken place so far.

We supply several internet based sets of Fama/French factors and are aware of five other providers of monthly factor time series for Germany. All may be downloaded free of charge and include the three factors SMB (small minus big stocks), HML (high minus low book-to-market firms) and the proxy for the market portfolio, which together constitute the Fama/French model, plus MOM (momentum, the stocks' past performance), the additional factor suggested by Carhart.[4] This is a fortunate situation in many respects: the factor sets can be compared and their weaknesses can be identified, not only from a theoretical perspective but with respect to the actual numbers. In addition, users have a choice and can check the robustness of their results. The five other providers are (the latter three cover several countries):[5]

(1) Artmann/Finter/Kempf/Koch/Theissen (2012b), University of Cologne, University of Mannheim (this was the first freely available data set for Germany and it is "on the air" since 2011, to our knowledge);

---

[1] Possible weaknesses have only been discussed recently, by e.g. Cremers/Petajisto/Zitzewitz (2010), Chen/Novy-Marx/Zhang (2011) and Asness/Frazzini (2013). The data in French's library is updated and improved continuously. A major recent improvement is the introduction of "Research Factors" in addition to the "Benchmark Factors".

[2] E.g., Franceur (HEC Montreal) provides a data set for Canada, Ammann/Steiner (University of St. Gallen) for Switzerland, Brailsford/Gaunt/O'Brien (2012) for Australia.

[3] E.g., two factor sets exist for the UK, one by Nagel, one by Gregory.

[4] We do not consider factor data sets which are offered by commercial data providers and data sets which are only offered on a daily and not on a monthly basis. We are only aware of one other daily data set on Germany, which is available by the academic data provider Eurofidai.

[5] We refer to the suppliers' paper(s) and/or offered data sets, mostly simultaneously.

(2) Hanauer/Kaserer/Rapp (2013), Universities of Munich and Marburg;
(3) Schmidt/Schrimpf/von Arx/Wagner/Ziegler (2011), University of Zurich, ETH Zurich, Aarhus University;
(4) Marmi/Poma, Scuola Normale Superiore di Pisa;
(5) Frazzini, Stern School of Business, New York / AQR Capital Management.

Calculating Fama/French factors for a specific country seems to be a well-defined task, but there are many details that must be decided upon, and many data problems that need to be solved. The choices to be made may involve costly consequences and uncertain benefits. For example, does the use of a larger and more representative data set on individual stocks increase the explanatory or predictive power of the factor time series? Or, does an investment in the correctness of the underlying data improve this power, or, do weaknesses of the data set on individual stocks just wash out because of the large number of stocks that enter the calculation on a specific date?

The objectives of this paper are to

- discuss the relevant issues and provide answers when possible;
- point out the major characteristics of the factor time series available for Germany;
- compare these time series empirically;
- give advice to factor providers and users.

Because the time series are used for different purposes and in different contexts we cannot answer the rhetorical question which is best, we only can point out strengths and weaknesses that may be crucial in specific applications.

Due to the explicit or implicit 'detail' decisions by the factor providers or the supplier of the underlying stock market database(s), the factor sets differ by (1) the included stock exchanges (about eight, their importance varied over time), (2) the included exchange segments (about three, with different stock characteristics), (3) whether and how they deviate from the calculation procedure suggested by Fama/French, (4) their treatment of German stock market peculiarities (e.g. dual class firms), (5) the underlying data bases, (6) whether the proxy for market portfolio is a publically available index or is calculated on the basis of the underlying data set and (7) by the covered time period.

We briefly discuss these differences individually.

At present, we only have reliable, double-checked data on the Frankfurt Stock Exchange (FSE), so we cannot include the other German exchanges (see Section 2.2). We strongly believe that the lowest market segment should not be included. Whether only the top segment or also the middle segment and the failed experiment Neuer Markt should be included is debatable (see Section 2.3). As a consequence we offer alternative time series, which differ in this respect. Artmann et al. and Hanauer et al. also concentrate on Frankfurt and exclude its lowest segment. The other providers do not differentiate with respect to stock exchanges and/or segments (see Section 2.2, 2.3 and Table 1).

The calculation procedures of all included factor sets closely follow Fama/French (1993). All providers use value-weight returns on six double-sorted portfolios. The factor returns are the differences of the equally weighted returns on subsets of the six portfolios, which means, that they may be interpreted as the results of long-short investment strategies (see Appendix F for factor calculations). However, there are two main areas in which possibly important deviations from the calculation procedures suggested by Fama/French exist.

First, Fama/French's proposal of using New York Stock Exchange (NYSE) breakpoints is typically not transferred to the German setting. They suggest to use the median market capitalization of all NYSE stocks to allocate the large amount of small stocks from other exchanges to the two portfolios "Small" and "Big", instead of using the median market capitalization of all included stocks. An alternative is to use a size breakpoint which is higher than 0.5. The suppliers of factor sets for Germany in general form the size portfolios using the median of all included stocks (that is, breakpoint 0.5). The only exceptions are Schmidt et al. (they offer two alternative sets, one with 0.5 and one with a size-breakpoint of 0.8), Frazzini (only 0.8) and our sets. We strongly believe in Fama/French's proposal but do not want to impose our beliefs on our users and provide alternative factor series (see Section 2.4).

The second deviation from the procedures of Fama/French is the exclusion of financials in several factor sets. Minor deviations from the calculation procedures of Fama/French and/or the other providers are in the areas of penny stocks and IPOs (see Section 2.8).

The fourth aspect in which the factor sets differ is their treatment of the German stock market peculiarities. German preferred stocks are economically very similar to common stocks. They therefore should be taken into account directly, by including them in a precise or approximate manner, or indirectly, by adjusting the book value of equity (see Section 2.5). Possibly some suppliers of international data sets do not take them into account at all. Another German peculiarity is the corporate income tax credit ("*Körperschaftsteuergutschrift*"), which existed between 1977 and 2000 and translated roughly to a 25 % increase of the cash dividend. We advocate that this tax credit is included in analyses of German stocks or mutual funds (see Section 2.6).[6] On our website we provide time series which include this financial benefit of stock ownership. In this paper we do not discuss these additional factor sets.

---

[6] It is, e.g., included in the database of the *Bundesverband Investment und Asset Management e. V.* (BVI).

The fifth aspect is the underlying data bases (see Section 2.1). Our factor sets, and those of Artmann et al. are based on academic stock market data bases that have been checked carefully and used in many prior studies. The data bases practically include all stocks listed on the chosen exchanges and segments during the covered time periods. Hanauer et al. only use Datastream in a time period in which good quality is documented, Schmidt et al. also use it for time periods in which quality has been found to be insufficient (Brückner (2013)). Marmi/Poma and Frazzini use commercial stock market data bases (Factset, XpressFeed Global) whose quality is not documented. In all three commercial data bases, the coverage of the stocks listed on a specific exchange or segment may not be documented in sufficient detail but improve over time. Occasionally, in these commercial databases the number of included stocks increases in large steps at specific points in time. And while an underlying data set may contain some data on a large number of firms, not all of the required data for factor calculation may be available for all firms. Often the book value of the equity seems to be missing in the underlying data set. As a consequence,

- a factor time series may have different qualities in the sub-periods it covers;
- the number of underlying stocks differs considerably within and between the time series we look at, especially in the eighties and nineties;
- the allocation of stocks to portfolios in a specific data set may change over time, not only because stock characteristics change, but also because of the availability of data on other stocks.

A sixth aspect is, whether the factor sets use a publically available index as a proxy for the market portfolio (see Section 2.7). Only Artmann et al. do this, but actually this could be a good choice for several other providers. All others calculate the rate of return on the market portfolio from their own sample. Comparing their time series with a publically available index would be an easy check of the quality of their underlying databases. In case large differences exist, it may be wise to use the publically available index.

Lastly, the factor time series differ with respect to the time period they cover. We and Artmann et al. go back to the late fifties (see Table 1), Schmidt et al. start in 1984, Marmi/Poma start in 1988, Frazzini in 1990 and Hanauer et al. in 1996.

Taken together, we analyze the following nine factor sets for Germany:

- Our 'TOP', which includes only the top segment of the FSE;
- Our 'ALL', which includes the top and the middle segment of the FSE and its Neuer Markt, applying 0.5 breakpoints;
- Our 'ALL' [BP: TOP], which includes all three segments but uses breakpoints based on the top segment;

- Schmidt et al. ‘s two factor series with size breakpoints 0.5 and 0.8;
- Frazzini’s factor set with size breakpoint 0.8;
- The three factor series supplied by Artmann et al., Hanauer et al., and Marmi/Poma that all use size breakpoints of 0.5.

In addition to the nine factor sets there are other time series which we include in our discussion. Cremers et al. (2010) suggest to construct factors based on common and easily tradable size and style indices. Dyck/Lins/Pomorski (2013) use local MSCI indices to calculate local Fama/French factors in their study of the international mutual fund industry, Cuthbertson/Nitzsche (2013) do it in their study of the German equity mutual fund industry. We include a SMB and HML time series based on MSCI indices for Germany to check whether they can compete with factor sets based on data for individual stocks, even though they are constructed by different methods (see Appendix G for details). We also include the “MSCI Germany” as proxy for the German market portfolio.

On his website, French provides returns for three portfolios of German stocks sorted on book-to-market. The portfolios include a relatively small number of stocks before 2008. Fama and French (1998) use the portfolios as left-hand-side assets in their pricing regressions, but because the portfolios are not well diversified, they do not construct country-specific factors.[7] We follow their lead and do not use their book-to-market portfolios to construct a value/growth factor for Germany. We do, however, include the German market portfolio from French’s website in our tests (“local returns, all four data items not required”). In addition, we include the currently most popular proxy for the German market portfolio, the CDAX calculated by the Deutsche Börse AG in our comparisons. For the time period after its official start in 1993, this index is a near ideal proxy for the German market portfolio. For earlier time periods, the quality of this time series is debatable. Finally, we include the DAFOX, the most popular proxy for the German market portfolio for prior years.

Several other academics have calculated and used Fama/French factors for Germany or other countries, but have not made them available freely on the internet (e.g. Ang/Hodrick/Xing/Zhang (2009)). These are occasionally given to other academics on request (e.g. Dimopoulos/Wagner (2012) use the database of Ang et al. (2009)). Some of these factor sets are calculated practically in the same way and are based on the same stock market data base as a publically available data set.[8] One of our main conclusions is that the quality of a factor set, to a large extent, depends on the choice of exchanges and segments, on

[7] Bessler/Drobetz/Zimmermann (2009) and Waszczuk (2013) use the portfolios provided by French to construct a HML time series for Germany, Wheatley/Quach (2013) do it for Australia.

[8] Annaert/De Ceuster/Verstegen (2013) note that the European dataset of Schmidt et al. is very similar to theirs.

the calculation procedures and on the underlying stock market data base. So our conclusions, to some extent, are also relevant for factor sets which are not made publically available.

All factor providers for Germany and the additional time series we look at cover at least the time period 07/1996 to 12/2011. For this time period, we compare means and standard deviations of the monthly time series on the market portfolio, HML, SMB and MOM factors. The mean of the market factor of Hanauer et al. is nearly twice as high as that of Marmi/Poma and Frazzini. Considerable differences also exist between the mean SMB and HML factors. We also analyze pair-wise correlation coefficients, perform mean (rank) tests and finally look at the largest differences between the estimates for individual months. We find that all correlations between the factor time series of Marmi/Poma and their counterparts are extremely low. The approximation of the SMB and the HML factor by MSCI size and style indices, as well as the SMB and HML time series of Frazzini, show relatively low correlations (mostly well below 0.4) with the other supplier's time series.We show graphically, that the differences between the factor means decrease considerably over time. This we interpret as strong evidence for our conclusion, that the major cause of the differences is the quality of the underlying data bases.

Even if we only look at the four 'best' performing factor sets (Artmann et al., Hanauer et al., Schmidt et al., and our 'ALL' with breakpoints from the top segment [BP: TOP]), considerable differences exist between the factors in a large number of months. This could bias the results of studies based on these time series, especially if the samples are small. The largest spreads for a given month are:

- MOM, 10/2001: 30.36 % (Hanauer et al.: -40.71, our 'ALL' [BP: TOP]: -10.35);
- SMB, 11/1998: 11.69 % (Hanauer et al.: 4.41, our 'ALL' [BP: TOP]: -7.28);
- HML, 10/2003: 11.54 % (Artmann et al.: 1.93, Schmidt et al.: 13.47);
- Markt, 11/1998: 10.92 % (Artmann et al.: 10.60, Hanauer et al.: -0.32).

Most of the largest spreads occur between 1998 and 2003, when the *Neuer Markt* existed. Setting up a good stock market database for this time period is a challenging task.

While comparing factor means, correlation coefficients and absolute deviations are interesting steps of our analysis, the ultimate test of the factors' comparative quality is to what extent the choice of a factor set affects the result of an empirical analysis. In a first application of the factor sets, we therefore analyze the individual performance of 41 mutual funds that focus on German stocks and also the combined 'Fund of Funds' with nine 4-factor time series. When we look at the full time period of 1996 to 2011, we get roughly similar results for all data sets, except for Frazzini and Schmidt et al. (with size breakpoint 0.5). The average alpha is positive for these two factor sets and negative for all other time series. However, nearly all factor loadings of the Fund of Funds have the same sign. In 5-year sub-periods, the results differ strongly. E.g., in the 5-year sub-periods of 07/1996 to 06/2001 and 07/2006 to 08/2011, some factor sets produce positive average alphas, while others produce negative average alphas.

In a second application, we look at 16 size/book-to market portfolios and at 16 size/momentum portfolios. We regress the excess return of these portfolios on the different Fama/French factor sets and test the null hypothesis that all 16 alphas are jointly zero by employing the the Gibbons/Ross/Shanken (1989) test (GRS test). We find considerable differences in the rejection rates (p-values) for the GRS test, but very similar average adjusted R^2.

There are several reasons, why we think our Fama/French time series for Germany are a good basis for empirical studies on the German market. Most importantly, we strongly feel that the quality of the underlying stock market data is very high (see Section 2.1). This claim is supported by our comparisons of the eight time series of market portfolios in Section 3.2: Four of them (CDAX, MSCI, French and our 'ALL') are highly correlated (R^2>0.98) and have very similar means. We are the only provider in this group that offers a full set of factors. In addition, two of our team members calculated the risk factors independently, using different programming languages. Their first round of calculations produced time series, which were highly correlated, but not identical with respect to specific factor estimates. After checking the programming procedures for two weeks, both produced time series that are practically identical. Another advantage is that on our website we offer six different sets of time series for different purposes as well as additional benchmark portfolios, on a monthly and daily basis.

The paper proceeds as follows. We first discuss the peculiarities of the German stock market; describe how we treat them and how they are treated by other suppliers. Included in this session is a discussion of the different time series that may be used as a proxy for the market portfolio. In Section 3, we first look at the number of observations in the underlying portfolios, then examine the different factor time series according to their statistical characteristics and their pair-wise correlations. In the following Section 4, we apply the different data sets of German Fama/French factors to 41 mutual funds that invest in German stocks. In addition, this section includes an application of the factor time series to 4x4 size/book-to-market and 4x4 size/momentum portfolios to evaluate their ability to explain them. The paper concludes by a set of recommendations for creators and users of Fama/French factor time series, which we hope is not only useful in the German context but also in an international context.

2 Factor Calculation in the German Environment

To calculate Fama/French factors for Germany, a number of decisions have to be made. The stock market and balance sheet database(s) have to be selected. The stock exchanges and market segments to be included in the factor calculations must be chosen. These decisions are not independent of each other, since the available databases do not cover all exchanges and segments with equal quality through time. It also has to be decided, to what extent the underlying data should be checked, whether outliers should be removed, how the market portfolio should be approximated, how preferred stocks and penny stocks should be treated, whether financials and IPOs should be included and what breakpoints should be used. In each case, a cost-benefit analysis has to be made: instead of using complicated procedures, e.g., to include preferred stocks, these could just be omitted from the analysis if the effort for including them outweighs the benefits. Instead of searching for data errors, outliers could be removed..

2.1 Data Sources

In the 1970s and 1980s, several academic efforts were made to create computer-readable databases containing German stock market data. Two of these efforts are still ongoing: Richard Stehle started a database of the Frankfurt stock exchange beginning in 1954, which is the basis of our factor sets.[9] Herrmann Göppl started a database that covers all German stock exchanges and all market segments beginning in 1974. The official name of the latter is *Karlsruher Kapitalmarktdatenbank* (KKMDB).[10] An improved version of this database is used by Artmann et al. (see Table 1).

Data on the German stock market is also included in the databases offered by Thomson Reuters (Datastream), Bloomberg, Xpressfeed Global, Factset and possibly other commercial data providers, which for recent years typically cover all exchanges and all segments. During the first years in which they cover the German market, they usually only include a small number of shares with unknown origin, probably most are listed in the top segment. E.g., Datastream's coverage of the German market starts in 1973 with less than 200 stocks. In 1988, the number of German stocks in Datastream increases from 200 to 600, so for the end of 1988 nearly all German stocks are included (see Table 2). The strength and weaknesses of Datastream are discussed in general by Ince/Porter (2006) and specifically for the German

9 We are grateful to the following persons who helped in this process over the years, most importantly: Anette Hartmond, Wolf Bay, Ralf Sattler, Jürgen Warfsmann, Norman Gehrke, Rainer Huber, Jürgen Maier, Christian Wulff (our Christian W.), Olaf Erhardt, Stefan Daske, Anja Schulz, Ralph Körstein, Olaf Grewe, Imre Kiss, Sven Brüsewitz, Christoph Barnekow, Julia Mummelthey, Matthias Lorenz, David Bosch, Stefanie Ahrens (Otte), Oksana Pryshchepa, and Ioana Sima.

10 The KKMDB is described by Bühler/Göppl/Möller (1993) and Herrmann (1996).

market by Brückner (2013). Brückner (2013) concludes that Datastream covers the Frankfurt top segment well from 1990 onwards, both with respect to data accuracy and coverage.

All of these databases only contain the data needed to calculate rates of return on individual stocks (prices, dividends, split data) and the number of shares outstanding. This allows the calculation of the market value of the equity and the rate of return on portfolios. Most of them do not contain balance sheet data and therefore do not allow the calculation of book-to-market ratios. To do this, additional databases must be used. In Germany, the *Aachener Bilanzdatenbank*, Worldscope, Compustat and others offer book values of equity in various ways (see Appendix C.1). A specific stock can only be included in the HML factor calculation, if both, the correct market data and balance sheet data are available and can be matched. Worldscope, which is typically used in combination with Datastream, e.g., before 1996 has only balance sheet for one third to half of the stocks included in Datastream. Using another source for balance sheet data would require a common identification code, which does not always exist. Schmidt et al. and Hanauer et al. use Datastream in combination with Worldscope (see Table 1). Marmi/Poma use the commercial data provider Factset, Frazzini uses the CRSP tape for the U.S. and XpressFeed Global for his international data set.

The main advantage of the KKMDB over its competitors is that it includes all German exchanges and segments starting in 1974. Our database is restricted to the FSE and only covers the top segment (since 1953), the middle segment (since 1987) and the *Neuer Markt* during the time it existed. Datastream, on the other hand, contains the largest number of stocks since 1989. It does distinguish between exchanges but not between segments, which is one of its major weaknesses.

The main advantage of our database is that we compared our data for the FSE top segment for the years 1974 to 1995 with the data in KKMDB, item by item. We found a large number of discrepancies, which we checked with other data sources. We then either corrected our database or notified Karlsruhe.[11] In addition, Brückner (2013) compared our data on the top segment with Datastream (1974-2007), also item by item. In case of a difference, he also went to other data sources, identified the correct data and corrected our data, if necessary. Thus, we strongly feel that with respect to the top segment, our data quality is better than Datastream, and KKMDB, the data bases typically used in the other studies (see Section C.5 for details about data quality). The very high correlation of our market portfolio with the CDAX, the MSCI and the time series of French, the very similar mean excess returns on these four time

[11] Due to the objectives and limited resources of the KKMDB, data errors (typos, missing data), to our knowledge, are not corrected in this database. Thus, each user has to check the KKMDB data by herself. This may have contributed to the fact, that the broad coverage of exchanges and segments by KKMDB so far has not been utilized by many users. From 1990 to approximately 2000, the KKMDB was the most important database for German academics. In recent years, to our knowledge, it has not been used by many researchers.

series and the near identical estimates of the SMB, HML and MOM factor means with those of Hanauer et al. support our confidence in the quality of our database.

The broader coverage of the German stock market by KKMDB and Datastream results in a much higher number of small companies in these databases compared to ours. However, this will only have a limited impact on the factor series if breakpoints from the top segments are used or if the size breakpoint is not 0.5 but 0.8 or even higher.

[Table 1]

2.2 Relevant Details on German Stock Exchanges

Since the 1950s, eight major stock exchanges existed in Germany. Düsseldorf, the center of the coal and steel industry, and Frankfurt, the centre of the banking industry, had equally important exchanges. Berlin, Munich and Hamburg also had important ones. The exchanges at Bremen, Hannover and Stuttgart were the smallest ones. In the fifties and sixties, most stocks were traded simultaneously at more than one exchange. However, only few were traded on all exchanges, and even fewer in the top segments of all exchanges (18 in 01/1958). Typically, the trading volume was highest at the *Heimatbörse* (home exchange). Over time, Frankfurt became more and more important, but in 1998 still only 540 of 883 (61 %) stocks were listed in Frankfurt. In 2009, 1025 of 1150 stocks were listed in Frankfurt (89 %).[12] Presently, Frankfurt is by far the most important stock exchange; the other German exchanges are often labeled as regional exchanges (*Regionalbörsen*).

Four factor providers specifically state, that they only include stocks traded in Frankfurt (we, Hanauer et al., Artmann et al. and Marmi/Poma), see Table 1. Our main reason for concentrating on Frankfurt is a cost-benefit consideration: Including other exchanges into our database would be very costly for the years before 1974, when digital data first became available in Germany. Most of the stocks not traded in Frankfurt are small stocks, for which it is difficult to get balance sheet data. Not including them may create a biased sample of the over-all German market, especially in the years up to the nineties. Using breakpoints from the top segment probably reduces this bias.

Frazzini does not mention the stock exchanges that are included in his data set. Schmidt et al. include all domestic stocks that are in Datastream, that is, all stocks from all German

---

[12] The numbers include all segments existing at the time (see section 3.1 and Table 2 for details, source: DAI Factbooks).

exchanges. This decision may be due to the fact that Datastream only includes the current exchange listing.

2.3 Market Segments

In addition to the 'horizontal' market segmentation, a 'vertical' one exists in Germany. During the time period covered by our study, three segments existed at most exchanges for most of the time. Because their names changed over time (see Appendix A), we will refer only to the top, the middle and the lowest segment.

On a given day, a specific stock may be listed in only one segment of a specific exchange. It may be listed in a higher segment at one or several exchanges and at lower segments at other exchanges. A specific stock may move to a higher segment of an exchange at any time, if it fulfills its listing requirement. In the long run, successful stocks typically move upwards and are listed at additional exchanges; unsuccessful ones move downwards and give up their listings at some exchanges. As a consequence, statistics about the number of listings are not very detailed before 1998 (see Table 2). It may also be difficult to obtain a complete history of a stock's market segment membership. As a result, a stock may be considered to be included in the factor calculations even though it is not in the defined universe of stocks.

For the top segments, a national law existed since 1896, so their regulation was nearly the same for all exchanges. The middle and the lowest segments were exchange regulated only. The regulation of the lowest segments was traditionally the weakest. Investor protection is still insufficient today: issuers are not required to publish a prospectus, firms are not obliged to publish ad-hoc announcements. A stock may be listed in the lowest segment without a formal application by the issuing firm, it suffices if a curb broker (*Freimakler*) beliefs that there is a demand for trading. The German Federal Financial Supervisory Authority (BaFin) classifies the lowest segment not as a stock market, but as a special form of a multilateral trading system[13] and has repeatedly warned investors about the special dangers associated with stocks listed in the lowest segment.[14] In addition, many press reports document market manipulations and other activities that may be labeled criminal. Criminal activity in the lowest segment has probably varied considerably over time. For instance on 14 April 2012, the Deutsche Börse AG announced that the *First Quotation Board*, which is part of the lowest segment, would be closed on 15 December 2012, because of multiple cases of market manipulation.[15] As a result, 96 German stocks (and many more foreign stocks) lost their FSE

---

[13] E.g. in the annual report 2011, p. 198.

[14] E.g. in their annual reports 2008 (p. 156), 2009 (p. 174), 2010 (p. 46), 2011 (p. 197), 2012 (p. 176).

[15] The *Frankfurter Allgemeine Zeitung* from 16 February 2012 describes the enormous extent of the criminal activities in this market segment in more detail, see http://www.faz.net/aktuell/finanzen/aktien/first-quotation-board-boerse-schliesst-marktsegment-11640303.html.

listing, which possibly was their only one. 16 stocks were already suspended one year earlier.[16] Taken together, the number of German stocks traded in the lowest segment in Frankfurt was reduced by more than 50 %.

The inclusion of stocks from the lowest segment may have considerable implications for the interpretation of the factor time series. To elaborate, until the end of the last century, the number of German stocks traded in the lowest segment of at least one stock exchange was small: In 1987, 147 out of 679 (22 %); in 1996, 123 out of 802 (15 %). In recent years, the number of German stocks traded in the lowest segments has increased drastically, in Frankfurt from 88 in 1998 to 424 in 2012 (see Table 2). The total number of stocks traded in Frankfurt in 2012 is 962, so 44 % of all stocks are traded in the lowest segment. Most of these stocks have a very low market capitalization and thus will be allocated to the portfolio of "small stocks" in the calculation of the factor SMB. Consequently, when half of the stocks are allocated to the small stock portfolio, this factor in recent years measures mainly the difference between the lowest segment and the two higher segments, which may be interpreted as a proxy for regulatory strength. For these reasons, we do not recommend including stocks listed in the lowest segment.

The AIM in London, in recent years, is in many ways similar to the lowest segment of the FSE. According to Gregory et al. (2011, p. 8), 2,087 UK companies were traded on the London Main Market in December 1998, only 307 on the AIM. By December 2008, the number of companies in the Main Market has fallen to 1,142; the number in the AIM has risen to 1,512 firms, of which 1,136 have market capitalizations of less than £25m.

All factor providers except Schmidt et al., Frazzini and Marmi/Poma specifically state, that they do not include stocks traded in the lowest segment, see Table 1. By looking at the number of stocks included by Schmidt et al. (see Table 2), we get the impression that they include a large amount of stocks of the lowest segment. This decision is probably related to Datastream's classification problem: i.e. based on Datastream data for the German market, one cannot distinguish between stocks from the different market segments. Since Datastream's coverage of German stocks increases through time, especially for the stocks from the lowest segments (Brückner (2013)), their sample may involve less stocks from the lowest segment in the beginning, and more in the end.[17]

While the three suppliers who provide factors for several countries (Schmidt et al., Frazzini and Marmi/Poma) seem to follow their data provider in the decision which segments to include, we, Artmann et al. and Hanauer et al. made this decision ourselves and collected the

---

[16] Souce: *Hoppenstedt Aktienführer* 2014.

[17] In addition, we noticed that Datastream's data quality and coverage improves over time. This makes it necessary for Schmidt et al. to redo the data collection, cleaning and factor calculation. Consequently, their Fama/French factor time series would continuously changing (improving).

additional data, if necessary, ourselves. Hanauer et al. use the stocks included in the CDAX to calculate the factors. This, however, implies, that before 1998 only the Frankfurt top segment was included, after 1998 also the middle segment and the *Neuer Markt*, which doubles the sample size between 1997 and 2000. Thus, the samples of Hanauer et al. and Schmidt et al. vary over time with respect to the included segments and the sample size. We analyze this in greater detail in Section 3.1.

Between 1987 and 2007, the middle segment was the *Geregelter Markt*. In their detailed analysis of this segment, Brückner/Stehle (2012) conclude, that this segment was regulated nearly as strictly as the top segment. We include this segment in four of our six factor series. Artmann et al. include the stocks of this segment only, if they were part of the top segment or the *Neuer Markt* at a later or earlier time, which we think may create biases.[18]

In the nineties, additional segments were introduced at some exchanges, most importantly the *Neuer Markt* in Frankfurt, which was opened on 10 March 1997. This segment was also exchange regulated and designed to give young technology based companies access to capital. The *Neuer Markt* may be compared to the NASDAQ in New York, the AIM in London, or the Nouveau Marché in Paris. This segment attracted a large number of IPOs of young technology firms and was therefore initially considered a tremendous success by many market observers, see e.g. Vitols (2001) and Kiss/Stehle (2002). However, many irregularities and the disastrous performance from 2000 to 2002 (burst of the dot-com bubble) severely damaged the reputation of this segment. The NemaxAllShare reflects the initial success and the catastrophic end of the *Neuer Markt*: Between 03/1997 and 03/2000, this index increased from 500 to 8559 points. From this peak, the index fell to its minimum of 358 points on 12 March 2003.[19] As a consequence, it was closed in 2003 and most of its firms were transferred to the middle segment.

The stocks listed in the *Neuer Markt* are included in all factor time series offered by other providers. We include these stocks in four of our six factor series. Two factor time series contain only stocks from the top segment of the FSE. We distinguish between the two data sets (pros and cons of the two data sets can be found in Appendix D):

I. 'TOP' – contains all firms, which have at least one class of shares listed in the top segment (*Amtlicher Markt*) of the FSE from 1953 to 2007 and, from 2007 onwards all

---

[18] Brückner/Stehle (2012) and Rasch (1994) indicate that firms that made it from the middle (*Geregelter Markt*) to the top segment (*Amtlicher Markt*) were among the most successful firms ("winners"), whereas average or low performing firms remained in the middle segment ("losers"). Also, non-top segment firms are usually relatively small compared to those in the top segment. Thus, the sampling scheme of Artmann et al. may introduce an ex post selection bias as well as a survivorship bias.

[19] Great care must be used in a standard risk/return analysis of this segment: From the beginning to the end of the segment, the arithmetic mean of the monthly returns was 1.07% %. The monthly geometric mean (compound return) was -0.31 %.

stocks in the *Regulated Market* that were formerly listed on the *Amtlicher Markt*. In this set, we do not include firms listed in the middle and lowest segment of the FSE and its Neuer Market or solely on other stock exchanges.

II. 'ALL' – contains all firms, which have at least one class of shares listed on the top segment (1953 to 2007), middle segment (*Geregelter Markt,* 1987 to 2007), *Neuer Markt* (1997 to 2003) and, from 2007 onwards all German stocks listed on the *Regulated Market* of the FSE. We do not include firms solely listed on other stock exchanges. We also do not include firms that are listed in the lowest segment of the FSE (its name changed several times, see Appendix A). Presently the time series also does not include the middle segment between 1953 and 1987 (named *Geregelter Freiverkehr*). This segment was relatively small; Brückner/Stehle (2012) discuss its history.

2.4 Breakpoints and Sample Selection for the Factor Portfolios

The SMB factor is calculated by subtracting the rate of return on the portfolio of big stocks from the return of the portfolio of small stocks. An important question is how individual stocks are allocated to these two portfolios. A simple solution is to rank stocks according to their market capitalization and put the top 50 % in the big portfolio, the bottom 50 % in the small portfolio. Fama/French (1992) observe that in view of the large number of small stocks, especially on the AMEX and the NASDAQ, the small stock portfolio would be dominated by the stocks from these exchanges. They suggest to use the median market capitalization of the NYSE to allocate the stocks from all three exchanges to the two portfolios "Small" and "Big". As a consequence, the portfolio "Small" contains much more stocks than the portfolio "Big",[20] but still only about 8 % of the combined market value of the two size portfolios.[21] The use of NYSE breakpoints has become a standard procedure in empirical studies focusing on the U.S. market.

In the UK, a "natural" set of breakpoints of this type does not exist. Gregory et al. (2011, p. 5) state in this context: "[O]ur central problem in forming the factors and portfolios is to find a UK proxy for the NYSE break points […]. This is an important issue as the London Stock Exchange exhibits a large 'tail' of small and illiquid stocks, which are almost certainly not part of the tradable universe of the major institutional investors that make up a large part of the UK market".

---

[20] In 1994, e.g., the portfolio of "small" firms of French's "U.S. Research Returns Data" (six portfolios formed on size and book-to-market) contains five times as many stocks.

[21] Fama/French (2012) use the market cap of a region to allocate stocks to the big and small portfolios. Small stocks are those in the bottom 10 % of total market capitalization.

In the German context, breakpoints based on the top segment could be used to allocate stocks from the lowest segments and the *Neuer Markt* to portfolios. We think this is a major issue and therefore offer two versions of the factor time series that include all Frankfurt stocks except those from the Open Market: One series that allocates the stocks according to the "50 % rule" and one that allocates the stocks from the lowest segments according to breakpoints based on the top segment. Schmidt et al. discuss the implications for the construction of the SMB factor as well and provide Fama/French factors using alternative size breakpoints of 0.5 and 0.8. Frazzini, only uses a size breakpoint of 0.8 in his international data set. Hanauer et al. and Artmann et al. only use the 50 % rule.

Another decision that a factor provider has to make is, whether the same or different samples of stocks should be used for all factor portfolios. If the same sample is used, a stock only enters the factor calculations if it fulfills the requirement to be included in all factors. The alternative is, to use different samples of stocks for the factors. In this case, only the requirements for a specific factor calculation need to be fulfilled.

All German factor providers discussed in this paper use a separate, typically a much larger sample of stocks for the calculation of the market portfolio than for the factor portfolios. An exception is Hanauer et al. who use the same set of stocks for the calculation of factors as for the market portfolio. Also, all factor providers for Germany, as well as French for the U.S., use identical samples for the HML and the SMB factor. Considerably differences exist with respect to the sample selection for the momentum factor MOM. We, Artmann et al. and Hanauer et al. use identical samples for SMB, HML and MOM. On the other hand, Schmidt et al. and Marmi/Poma use a separate, mostly larger sample for the calculation of the momentum factor, than for SMB and HML factor. The choice of Frazzini is unknown to us.

2.5 Dual Class Firms

Many German firms issue two classes of stocks, *Stammaktien* (typically translated with common stocks) and *Vorzugsaktien* (non-voting stocks that are typically translated with preferred stocks). The risk-return characteristics of German preferred stocks are very similar to those of common stocks, while U.S. preferred stocks are very similar to bonds.[22] German preferred stocks in fact are very similar to the U.S. common stock class of dual-class firms, which have inferior voting power.

The fraction of firms in the top segment of the FSE that issues preferred stock increases from 8 % in the late fifties to 15 % around the end of the century, and then decreases to 8 % in 2007. Typically, both classes of stocks are traded on an exchange, preferred stocks typically

[22] In the U.S., dividends on preferred stocks typically have a specified level and must be paid before common stockholders receive dividends.

with a discount.[23] Appendix B describes the major differences between German preferred stock and U.S. common stocks with inferior voting power and dual class firms in the U.S. and Germany in greater detail.

As a consequence of the existence of two share classes with equity like characteristics, the way in which they are included in the analysis may be an important decision. The following solutions are feasible:

- Common and preferred stocks could be treated as a unit, that is, firm observations enter the analysis. The rate of return could be calculated either by just using one of the two types of stock prices or by using both prices.
- Common and preferred could be included as separate observations. In this case, either the book value has to be split before the book-to-market ratios are calculated or the same book-to market is used for both.
- Only one stock type is included, typically the more important one. In this case, the book value has to be adjusted.

We use the first procedure, that is, we calculate the rate of return on the total equity portfolio of a firm. Whenever both classes are exchange listed, we calculate this rate of return precisely. When one class of shares is not listed, we use the prices of the listed type to estimate the rate of return on the firm's total equity. We estimate the (total) market value of a firm's equity by aggregating the market value over all share classes (details can be found in appendix C.2). Most researchers simplify this procedure by using the rate of return of only one class as an estimate of the rate of return on the firm's total equity. Typically, the rate of return on the common stocks is used. Artmann et al. use the class of stocks for which the longer return time series is available. However, they also sum up the market value of the share classes to obtain the firm size. Hanauer et al. use firm observations too (see Table 1). On the other hand, Schmidt et al., Frazzini and Marmi/Poma remove non-voting shares and thus possibly obtain upward biased estimates of the book-to-market ratio of dual class firms.

2.6 Tax Imputation System (Körperschaftsteuergutschrift)

Prior to 1977 dividends were taxed at both, the corporate and the personal level. From 1977 to 2000, the double taxation of dividends was eliminated for German shareholders. In addition to their 'cash dividend', they received a voucher from the tax authorities in the amount of the corporate income tax that was paid on their dividends (corporate income tax credit*, Körperschaftsteuergutschrift*). This voucher could be used to pay the personal income tax or to receive a tax refund.[24] From 1977 to 1993, the value of these vouchers was 9/16 (56.25 %) of

---

[23] Daske/Ehrhardt (2002) describe the price and return characteristics in detail.
[24] See Stehle/Hartmond (1991) or Murphy/Schlag (1999) for more details about tax credits.

the cash dividend. As a consequence of the reduction of the corporate income tax rate to 30 % it was 3/7 (approx. 42.86 %) of the cash dividend from 1994 to 2000. This "imputation system" ended in 2001.

On average, large German firms pay higher dividend yields than small firms. Therefore, not including the tax credit for the corporate income tax has practically the same effect as not including the dividends in the rate of return calculation. On average, the rates of return are biased downwards; the bias is greater for large firms than for small firms. This would weaken a potential size related effect or increase a reverse size related effect.

We therefore provide both data with and without taking the corporate income tax credit into account. This allows researchers to use the proper data set for their analyses. None of the other factor suppliers we include in our analyses discusses, nor takes the corporate income tax credit into account. If the corporate income tax credit is not included in the calculation of the rates of return of individual stocks or in the indices, the calculated return is equal to the after-tax return of an investor with a marginal tax rate of 36 % (30 % after 1994).

2.7 The Rate of Return on the Market Portfolio

An important part of the service offered by providers of Fama/French factors is that they include a time series on the rate of return on the market portfolio. This is important because the available index data for the German market has many weaknesses that must be taken into account in an empirical study.

Presently, the most prominent proxy for the German market portfolio is the CDAX (performance index) offered by the Deutsche Börse AG. The CDAX officially started on 22 April 1993 and includes until 21 September 1998 only stocks traded in the top segment of the FSE (*Amtlicher Markt*). Since then it also includes the stocks listed in the second segment of the FSE (*Geregelter Markt*) and from 1998 to 2003 also the stocks listed in the *Neuer Markt*. Thus, the CDAX is a good proxy for the performance of the stocks traded in the top segment of the FSE between 1993 and 1998. After 1998, it is a good proxy for the performance of all stocks of the FSE with the exception of the stocks traded in the lowest segment.[25]

The basic concept of the CDAX index is the same as that of the DAX, which is very good except that it does not include the corporate income tax credit (see section 2.6). Another problem, which we already mentioned, is that it changed the segment coverage in 1998. The Deutsche Börse AG also provides a CDAX time series that covers the period from 30

[25] The DAX index family only includes stocks that are traded on the FSE. Indices that cover or include other exchanges, e.g. the *Aktienindex der Westdeutschen Landesbank* have not played an important role after 1990. To our knowledge, none of the stocks that are not listed in Frankfurt are included in any index or benchmark that is used in empirical studies on the German market.

December 1987 until 22 April 1993. To our knowledge, its composition is not documented and may contain, for example, an ex-post selection bias. From 1970 to 1988, the 'official' CDAX is based on the *Frankfurter Wertpapierbörse Index* (*FWB-Index*). The *FWB-Index*, like most indices at the time, did not take dividends into account.[26]

Many studies on the German capital market that focus on time periods between 1974 and 2004 are based on the *Deutscher Aktienforschungsindex* (DAFOX). This time series is available from 1960 to 2004 and only covers the top segment of the FSE.[27] From 1960 to 1974, only fractions of the stocks of this segment are included. To us, the DAFOX return in 1998 is way too high and thus most likely incorrect: it is 32.9 % whereas it is much lower in all other time series (CDAX 15.5 %, our 'TOP' 17.9 %, our 'ALL' 18.7 %, Schmidt et al. 17.5 %, Hanauer et al. 14 %, French 18.4 %, Frazzini 17.4 %, MSCI 20.3 %). From 1974 to 1996, the DAFOX is practically identical to our 'TOP' market return. Before 1974 the differences between the two series are somewhat larger. In addition, the DAFOX does not include the corporate income tax credit.

Another proxy for the German market portfolio is the MSCI Germany Index. We also include this index in our analyses since it is used in many studies that cover a large number of countries. This index commences in February 1970. It is only based on 55 large and mid cap stocks.

Because of these and other problems, many empirical asset pricing studies (e.g. Schulz/Stehle (2002), Schrimpf/Schröder/Stehle (2007), and Ziegler/Eberts/Schröder/Schulz/Stehle (2007)) calculate a proxy for the market portfolio from their own sample. All factor suppliers follow this approach, except Artmann et al., who use the DAFOX (until 2004) and the CDAX (starting in 2005). Thus, until 2004 their market portfolio consists only of the stocks from the top segment whereas their HML, SMB and MOM factors contain all firms listed in the top segment, the Neuer Markt and from the middle segment if they were part of the top segment or the *Neuer Markt* at a later or earlier time (see Table 1 and Section 2.3). To users who need a time series going back to the seventies, we therefore recommend not to use the market portfolio of Artmann et al. but to use our 'TOP' or 'ALL' time series or, to replace the

[26] The CDAX time series is easily accessible on the website of the Deutsche Bundesbank and is also included in Datastream. On the Bundesbank website, the CDAX price index has a higher rate of change than the CDAX performance index in many years before 1988. This should not happen, since the only difference between both indices should result from a different treatment of dividends. In 1985, e.g., the price index changes by 71.9 %, while the performance index only by 56.9 %. The CDAX performance and price index in Datastream are consistent in this respect. Thus, before 1988, we do not recommend to use the CDAX performance index from the Bundesbank website.

[27] The DAFOX is documented in Göppl/Schütz (1995) and available from the KKMDB. Among others, Schlag/Wohlschieß (1997), Wallmeier (2000), Elsas/El-Shaer/Theissen (2003), and Artmann et al. use the DAFOX.

DAFOX by the CDAX already after 1993. Following this advice, a user also avoids the DAFOX error in the year 1998.

2.8 Other aspects

In Fama/French (1992), financial firms (mainly banks and insurance companies) are specifically excluded from the portfolios that are used as test assets. In this context, they argue (p. 429): "We exclude financial firms because the high leverage that is normal for these firms probably does not have the same meaning as for nonfinancial firms, where high leverage more likely indicates distress." This argument refers to their inclusion of leverage variables in the set of independent variables. Since there is no leverage factor in the 1993 paper, financials are included in their factor calculation. The large majority of papers on the cross section of returns and on factor models, focusing on the U.S. market, do not exclude financials from the test asset portfolios. However, a number of pre 1998 papers do exclude them (e.g., Jagannathan/Wang (1996), Knez/Ready (1997) and Loughran (1997)). Barber/Lyon (1997) show that the relationship between security returns, firm size, and the book-to-market ratio is similar for financial and nonfinancial firms. All portfolios on French's website currently include financials, also the two providers of British factor time series.

These subtle differences are not discussed by most factor providers. Only Artmann et al. and Hanauer et al. mention this question. Several earlier studies on the German market exclude financials (e.g., Schulz/Stehle (2002), Schrimpf et al. (2007), and Ziegler et al. (2007)) and/or point to the special characteristics of financials. However, in view of the results of Barber/Lyon (1997) and the treatment of financials by Fama/French, including financials may be preferable. On the other hand, Viale/Kolari/Fraser (2009) find no evidence that bank stocks have an exposure to size and book-to-market. So the exclusion of financials from the factor portfolios may not matter much. We exclude them because of data limitations. Artmann et al., Hanauer et al. and Marmi/Poma also exclude them in the calculation of SMB, HML and MOM. Hanauer et al. even exclude them from their market portfolio. Schmidt et al. and Frazzini include financials in all calculations.

Some stock exchanges, for example the NASDAQ, delist penny stocks, stocks with a share price of less than one.[28] In contrast, in Germany stocks are not delisted according to such market rules.[29] We observe that most penny stocks are stocks of (nearly) bankrupt firms that no longer publish financial statements. The rate of return on penny stocks typically has a

---

[28] See NASDAQ Stock Market Rules, Rule 4000 Marketplace Rules, The Bid Price Requirement, http://cchwallstreet.com/nasdaq. The SEC refers to penny stocks as "low-priced (below $5), speculative securities of very small companies." See URL: http://www.sec.gov/answers/penny.htm, 12 August 2011.

[29] In 2001, the Deutsche Börse AG attempted to delist penny stocks from the *Neuer Markt*, but was not successful, see Brückner/Stehle (2012).

much higher standard deviation than the rate of return on stocks with higher prices, because minor price changes might yield rates of return of 100 % or more. The German BaFin regularly warns German investors that price manipulations occur frequently in the penny stock area.

Artmann et al. and Hanauer et al. keep penny stocks in their sample. On the other hand, Schmidt et al. delete stocks "with prices less than one unit of the domestic currency". This procedure excludes stocks like Infineon – they traded below 1.00 EUR in early 2009 – which in fact is a member of the DAX with a considerable market capitalization. We therefore classify stocks whose share price is below one and whose market value of all equities (ME) of the firm is less than five mln. as penny stocks and exclude them. Our data sets contains penny stocks according to this definition and only from July 2001 onwards (data set 'ALL' has 2001:4, 2002:40, 2003:32, 2004:16, 2005:10, 2006:2, 2007:5, 2008:20, 2009:18, 2010:12, 2011:8, 2012:7).

We also remove profit participation bonds (*Genussscheine*) from our data set, which are also not considered in our estimates of the book values of equity due to their debt character (Stehle (1994)). Finally, we remove *Restquoten*, from our data set, a common practice in German papers. IPOs and firms listed for the first time are added to our data set at the end of the month of their first listing. However, since we require, among others, a market value of equity as of December of the previous year and a return history of at least 12 month, IPOs are effectively included in the portfolios at earliest after 12 month, at latest after 17 month. An exception is the time series for the market portfolio, for which we only require a return and market value of equity as of the previous month (see Appendix E).

3 Comparison of the Fama/French Time Series for Germany

3.1 Number of Observations in the Underlying Portfolios

The suppliers of Fama/French factor time series we analyze differ considerably with respect to the treatment of the peculiarities of the German stock markets that we have discussed in Section 2. In addition, we argue, that the individual coverage of the chosen market(s)/segment(s) is important. Basically, all stocks that are traded in a chosen segment on a specific date should enter all calculations for that date. If a few stocks are missing because of data limitations it may not matter much, if the remaining sample is still representative of the chosen segments. However, an inadequate or very unstable coverage may seriously undermine the quality of the calculated factors. We therefore compare the number of included observations in the 2x3 size/book-to-market and 2x3 size/momentum portfolios of the different factor suppliers. Table 2 shows these statistics for selected years as well as the number of stocks listed at different segments of the FSE.

[Table 2]

Over the years, the number of firms that enter our size/book-to-market and size/momentum portfolios is on average 28 % lower than the number of individual stocks that we include in the calculation of the rate of return on the market portfolio (e.g. in December 2004: 227 vs. 324 in the data set 'TOP', 480 vs. 691 in the data set 'ALL'). The difference is mainly caused by firms which issue different types of equity (dual class firms, see Section 2.5) and by the fact that we include financials in the market portfolio but not in the factor portfolios (see Appendix E for the calculation). Additional causes for differences in our time series are:

- Firms enter the market portfolio immediately after their first listing, typically their IPO, while they only enter the size/book-to-market portfolio with a delay of at least 12 months (12 month is the return history requirement, "at least", see Appendix E). This increases the differences considerably in 'hot' IPO years, especially in 1999 and 2000.
- The underlying data is insufficient, e.g. book values of the equity are missing.

The latter cause is a minor problem in our data set, since we hand-collect data if it is not available in the databases we use. We have the impression, however, that the latter cause is a more important reason in some of the other factor sets. For example, Schmidt et al. in the late 1980s, especially in 1988, have a market portfolio consisting of 406 observations, whereas their size/book-to-market portfolios contain only 142 observations (65 % missing). They use Worldscope as the only source for the book values of the equity. Possibly the increase in the number of stock observations in Datastream (1988: 285 to 662) is not matched by a similar increase in balance sheet data in Worldscope. This points to an important issue: Datastream's coverage of the German stock market increases over time, as well as its quality (see Brückner (2013)). As a result, we think that the quality of the factor time series by Schmidt et al. increases through time.

On the other hand, the number of observations in the market portfolio by Hanauer et al. is the same as in the size/book-to-market and size/momentum portfolios. They calculate the rate of return on the market portfolio from the same sub-sample of stocks for which they have the data for <u>all</u> other factors available. We think this is unusual and recommend users to replace the Hanauer et al. time series on the rate of return on the market portfolio with the CDAX, available from the Deutsche Bundesbank (see Section 2.7). Additionally, this adjustment eliminates the problem of not including financials in the market portfolio.

Marmi/Poma seems to have large coverage problems between 1993 and 1998. In 1993, there are 235 observations included in the size/book-to-market portfolios, two years later only 76. Similar issues can be observed for their size/momentum portfolio, which is based on 225 observations in 1993, but only on 72 for the next year. In addition, the size/momentum factor time series contains 16 gaps (NA values) in the middle of the nineties. Unfortunately, we do not have access to the number of stocks included in the calculation of their market portfolio. For the data set of Frazzini, we only have available the mean number of stocks included in his factor set. Asness/Frazzini (2013) document 662 stocks for the time period of 1989 to 2011, which is less then Schmidt et al. (797, market portfolio, same period) but more than our 'ALL' (561).

The data provided by French contains a very large increase in the number of included stocks in 2008. In 2007, the market portfolio is based on 89 stocks, in the following year on 1013. This increase seems to be substantial but has only little impact on the rates of return on his market portfolio. Most likely only small- and middle-sized stocks are added. The underlying data by Artmann et al. does not have these problems; their coverage is always good. In some months, especially in the early years, for which only we and Artmann et al. offer data, their number of observations included in the market portfolio as well as in the size/book-to-market and size/momentum portfolios is slightly larger than in our factors.

3.2 Characteristics of the Factor Time Series

This section compares the different Fama/French factor time series, mainly by examining their descriptive statistics, correlations and monthly differences.[30] We focus on the time period of 07/1996 to 12/2011 for which the data of all suppliers is overlapping.

3.2.1 Means and Standard Deviations, Stability over Time

At a first glance, the monthly mean excess rates of return on the market portfolio differ considerably, they vary between 0.28 % and 0.56 % (see the top panel of Table 3). However, six of the ten time series have relatively similar means: CDAX, MSCI, French, Schmidt et al., our 'TOP' and our 'ALL'. They are all between 0.39 and 0.46 % for the time period 07/1996 to 12/2011.[31] This implies a market risk premium of 5 to 6 % per year. The CDAX, Schmidt et al. and our 'ALL' estimate are at the lower end, 0.39 to 0.42 %, probably because more small stocks are included. The MSCI, French's time series and our 'TOP' concentrate on larger stocks. They have higher mean returns, which reflects the reverse size effect (0.44 to 0.46 %).

---

[30] We do not include a comparison of the risk free rates provided by the suppliers, as they are essentially identical. We always use our estimates of the risk free rate (see Appendix C.4) to calculate market excess returns (R_m – R_f).

[31] We use the CDAX performance index that is freely available on the website of the Deutsche Bundesbank. In the time period 1996 to 2011 this CDAX time series implies nearly identical rates of change as the CDAX in Datastream, the largest monthly difference is 1 %.

[Table 3]

The estimates of Marmi/Poma (0.31 %) and Frazzini (0.28 %) are much lower than the other means. Although, the difference is not statistically significant (see Table 4, middle and right-hand side column), it may crucially affect the outcome of specific types of studies, e.g. event-studies and mutual fund performance studies (see Section 4.1). Their low mean returns are possibly due to the fact that the underlying databases do not include all stock splits, stock dividends and rights issues. All three events make important contributions to the rate of return on German stocks.

We attribute Hanauer et al.'s high mean return (0.56 %) to the fact that they exclude financials from their market proxy since they did not perform well in the time period we look at. We attribute Artmann et al.'s high mean return (0.53 %) to the DAFOX error in 1998 and the non-inclusion of the *Neuer Markt*.[32] The standard deviations of the rates of return on the market portfolio of all providers are similar (except Marmi/Poma), also the kurtosis (again, except Marmi/Poma) and the skewness (again, except Marmi/Poma).

Figure 1 to 4 show rolling factor means, each observation is based on the prior 60 months. Figure 1 focuses on the monthly excess returns on the different proxies for the market portfolio. It shows that in the first 5-year period, which ended in 06/2001, the mean of the time series provided by Marmi/Poma differs from the other time series by at least 0.5%. This rolling mean is considerably lower than the others until 12/2011. The rolling 5-year means of the other time series differ by 0.4% in 06/2001, in the years ending after 01/2008 all except Hanauer et al. are nearly identical. Hanauer et al. do not include financials in the market portfolio.

[Figure 1, Figure 2, Figure 3 and Figure 4]

Figure 2 focuses on the SMB factor, Figure 3 on HML, Figure 4 on MOM. All show the same general pattern: In 2001-6 the differences between the rolling means are large, in 2011-12 the differences are much smaller. This we interpret as strong evidence for our conclusion, that the major cause of the differences is the quality of the underlying data bases.

In the SMB figure (Figure 2) the rolling means differ by 1.0 %, in 06/2001, in 12/2011 only by 0.5 %. The rolling mean of Frazzini is at the top at the beginning and at the end. Our ALL

[32] In their empirical test of the CAPM, Artmann et al. include Neuer Markt stocks in their left-hand side portfolios. This and the use of the DAFOX on the right-hand side may have caused their conclusion "none of the models [CAPM, 3-factor, 4-factor] can consistently explain the cross-section of [German] returns."

series is at the bottom at the beginning and at the end, closely accompanied by the very similar series of Hanauer et al. and Artmann et al.

In the HML figure (Figure 3) the rolling means differ by 1.5% at the beginning and at the end. Note however, that MSCI's and Frazzini's time series are 'outliers' at the end. The other means differ at the end by only 0.5 %.

Most impressive are the MOM rolling means. During the last three years they are nearly identical.

We conclude this section by additional observations on the over-all means and the rolling means.

The differences between the over-all means of the SMB, HML, and MOM time series are even larger than the differences between the proxies for the market portfolio. While it is easy to make an educated guess why the mean returns on the market portfolios differ, the differences between the factor time series are more difficult to explain, without underlying data sets. For all three factors the estimates of our 'ALL' series is very close to the estimates of Hanauer et al. This is not surprising, since the construction details, the market segments and the databases of both series are very similar.

Our two time series for the SMB factor, which include all segments but differ with respect to the breakpoints used (data sets "ALL") differ considerably from each other: the over-all means are -0.48 versus -0.80 % (see Table 3). This shows that the choice of breakpoints is an important issue (see Section 2.4). For the stocks from the top segment of the FSE ('TOP'), the mean SMB factor is even lower (-0.32 %). To a large extent, this is probably caused by the exclusion of the *Neuer Markt* stocks. Schmidt et al., who seem to include stocks from the lowest segment of the FSE (Open Market, see Section 2.3) and/or from other exchanges, provide two Fama/French factor sets, one using a size breakpoint of 0.8 and another with 0.5. Surprisingly, the means of the two SMB factor time series are very similar (-0.28 and -0.27 %) and are considerably lower than the other estimates. We would expect to see a similar difference as we obtain from our data when using different breakpoints.[33]

Other details of our SMB results we find interesting are:

- The over-all mean of the SMB time series of Artmann et al. It should be similar to our series 'ALL', but actually it is most similar to our series 'ALL' with breakpoints from

[33] Between 1996/07 to 12/2011, the mean number of stocks in the 'big' portfolio of Schmidt et al. [size breakpoint 0.8] and our data set ('ALL' with size breakpoint from the top segment) are very similar: 149 vs. 139. The mean of our 'big' portfolio is 0.863 %, the 'big' portfolio of Schmidt et al. amounts to 0.803 %. Additionally, the plots of these two big portfolios (not shown) are very similar. Thus, the difference in the SMB factors results from the 'small' portfolios. Additionally, when we use a 0.8 size breakpoint for our data set 'ALL', we get a very different mean (-0.42) for SMB than the one we obtained with breakpoint 0.5.

the ‘TOP” data set. One reason may be the different market segment selection of Artmann et al (see Section 2.3).

- The SMB factor approximation based on MSCI indices and Frazzini’s time series have by far the highest over-all means (-0.09 % and -0.03%).
- Judged by its over-all mean, the SMB factor of Marmi/Poma seems to be reasonably good (-0.56 %), but its skewness and kurtosis, which are far from other estimates, point to problems in this time series.
- The reverse size effect is large in the 5-year period ending 2001-06. It becomes smaller after 2002

Nevertheless, note that all factor providers agree on two important aspects:

- According to the means, a reverse size effect existed in Germany from 1996 to 2011.[34] This is different to most other capital markets, especially the U.S.
- The reverse size effect is most pronounced from 1996 to 2001.

The HML factor tries to capture that high book-to-market firms (value stocks) tend to perform better than low book-to-market firms (growth stocks). Again, Table 3 shows the descriptive statistics for the different data suppliers. The over-all means of the HML factors are, in absolute terms and on average, considerably higher than those of the SMB factors and nearly all are statistically different from zero. Again our results show, that the choice of breakpoints makes a difference. On a first glance, the mean return of the HML factor time series of our data set ‘TOP’ (0.42 %) look different to the others. Primarily, this is because this data set contains only stocks from the top segment that are potentially value stocks. Therefore, a value premium is less likely to be prevalent. The time series means of the two HML factors of Schmidt et al. are 0.55 % (size breakpoint 0.5) and 0.77 % (size breakpoint 0.8) also highlight that breakpoints can make a considerable difference (see Section 2.4 for a discussion). The estimate by Artmann et al. (0.85 %) is the second largest. The mean of Marmi/Poma’s HML factor time series (1.13 %), again, questions the reliability of their data.

Figure 3 shows that a value effect existed over the entire time period we look at, it was most pronounced from 1998 to 2005. The HML factors based on MSCI indices are in line with the others before 2007, but from then onwards are considerably lower.

The descriptive statistics for the momentum factor (MOM), demonstrate that buying recent winners and short-selling recent losers seems to be profitable, at least if transaction costs are not taken into account. All over-all mean returns on the MOM portfolio are well above 1 % per month (see Table 3), which is considerably higher than the over-all means of the other

34 See Amel-Zadeh (2011) and Stehle (1997).

factors. The lowest estimate is 1.05 %, by our data set ‘ALL’, the second lowest is Hanauer et al.’s (1.19 %). The highest mean MOM factor is provided by Marmi/Poma (1.60 %), the second highest by Artmann et al. (1.44 %).

However, all factor suppliers provide evidence that momentum seems to be very strong in German stock returns. This can also be seen in Figure 4, showing large rolling means. In comparison, for the same time period of 07/1996 to 12/2011, the mean of the momentum factor by French for U.S. stocks is only 0.47 %. Also, we note that the minima and maxima as well as the standard deviations of the means of the MOM factor time series are higher than for the other factors and the market portfolio. Finally, all presented MOM factors for Germany have a negative skewness, which is similar to the MOM factor in the U.S. (Daniel/Moskowitz (2011)).

3.2.2 Correlation Coefficients and Tests of Means

To identify similarities and differences between the factor sets, we also calculate correlation coefficients. Table 4 shows R^2s from simple OLS regressions where the left hand variable is a factor time series, explained by a time series of another supplier. To conserve space, we only include our time series ALL with breakpoints from the top segment and the Schmidt et al. series with a size-breakpoint of 0.8 in the analysis, even though the excluded time series ‘ALL’ is by construction more similar to the time series provided by others, notably by Hanauer et al.

[Table 4]

The correlations are highest between the different time series for the excess return on the market portfolio. In these time series large stocks dominate the results and the different breakpoint choices do not matter. The MSCI time series is very similar to the CDAX, the R^2 is 0.99, even though the CDAX contains a much larger number of securities. The series provided by French and our ‘ALL’ series have an R^2 with the CDAX of 0.98, next are Artmann et al. (0.97), Schmidt et al. (0.89) and Hanauer et al. (0.85). The lowest R^2 between any pair of the mentioned factor time series is 0.82. The four time series with the highest pair-wise correlation (CDAX, our ‘ALL’. MSCI and French) also have very similar means. They all include financials, focus on the top two segments of the FSE and include the *Neuer Markt* during its existence. If this is the correct universe of stocks that should be included in the market portfolio (from a theoretical perspective), then these four time series are probably the best proxies.

The next highest correlations are between the momentum time series: The lowest R^2 between any pair of Artmann et al, Hanauer et al., Schmidt et al. and our 'ALL' is 0.63, the highest is 0.83. A bit lower are the correlations between the SMB factor time series: The lowest R^2 between any pair of these time series is 0.51, the highest is 0.77. Much lower are the correlations between the HML factor time series: The lowest R^2 between any pair of the time series is 0.37, the highest is 0.49. We believe that the differences in R^2 between the factor groups are mainly caused by the quality of the underlying databases for the stocks. To calculate HML, book values of equities are needed; some of the underlying databases may be more complete than others or may have different values.

Table 4 contains p-values of t-tests and Wilcoxon tests to determine if the means or mean ranks of two sets of factor time series are significantly different from each other.[35] Both tests reject the null that the market excess return of Schmidt et al. and Hanauer et al., as well as Frazzini compared to Artmann et al. and Hanauer et al. are the same on a 10 % level. And both also reject that the SMB factors of Schmidt et al. and Hanauer et al. are the same. This may relate to the fact that the underlying samples differ in many aspects. For example, Schmidt et al. has a much larger sample and includes financials, while Hanauer et al. does not include financials into the market portfolio and the other factor portfolios. Also, both pair-wise tests reject the null that the MSCI based estimates for SMB and HML, and the SMB factor by Frazzini are the same compared to several others.

3.2.3 Economically Significant Differences

To identify systematic errors in the time series we compute significant differences between the normalized factor time series (mean zero, standard deviation one) and calculate absolute differences between single data points (months). We normalize the time series, because we want to compare the number of errors across factors. Two data points with an absolute difference larger than one are defined as a significant difference.[36] Again, to conserve space, we only include our time series 'ALL' with breakpoints from the top segment and the Schmidt et al. series with a size-breakpoint of 0.8 in the analysis.

Table 5 shows the results for the market portfolios as well as for the SMB, HML and MOM factors. Each part of Table 5 (e.g., the first part which contains the results for the excess rates on return on the market portfolio) is subdivided into two sections. The top one looks only at the differences within the group of the four 'major' providers that performed best in our

[35] Since the factor time series are related to each other, we assume dependency between them. We therefore perform dependent t-test for paired samples (assuming normality) and Wilcoxon signed-rank test (no normality assumption necessary).

[36] We also looked at the differences that are larger than two standard deviations. These are less but have the same general pattern.

preceding theoretical and empirical analysis (Artmann et al., Hanauer et al., Schmidt et al. with a size-breakpoint of 0.8, and our ‘ALL’ with breakpoints from the top segment). The bottom section includes the other providers.

[Table 5 and Table 6]

The results in Table 5 are fully in line with the results of our analyses of means and pair-wise correlations:

- In the time series for the market portfolios, relatively few large differences can be found, except for the Marmi/Poma market proxy. Within the four market portfolio proxies we consider to be best (our ‘ALL’ compared to MSCI, CDAX, French), only two economically significant differences exist.
- The number of differences in the MOM series is a bit higher.
- Somewhat even higher is the number of differences in the SMB factors.
- The differences are by far the largest in the HML time series.
- The factor approximations based on the MSCI size and style indices for Germany and the factor time series of Marmi/Poma and Frazzini deviate often from the other SMB and HML series.

Within the group of the four ‘major’ factor providers, the SMB factor time series by Schmidt et al. exhibits a relatively high number of significant differences (76, divided by 3, a mean of 25.33). Most of the differences are related to our factor series ‘ALL’ with breakpoints from the ‘TOP’ data set (26) and to the factor series of Artmann et al. (28). Also, the SMB factor of Schmidt et al. has eleven data points (month) where the time series deviates significantly from all three other ‘major’ suppliers. It might be related to financial firms that are included in the sample of Schmidt et al., but are excluded in the other three. Also worthwhile to note is that the results do not change considerable if we use the time series of Schmidt et al. with a size breakpoint of 0.5 (73 significant differences).

The numbers of significant differences between the HML factor time series are about two times higher than between the SMB series, but without an outstanding number for an individual series. The overall higher numbers of significant differences are most likely due to the fact that the book values of equity data may differ between the different data suppliers.

Finally, Table 6 looks at the months with the largest spreads between the (non-normalized) factors, it is based on Table 5. The absolute values of the spreads demonstrate that in applications with a small number of observations, e.g. event studies, the outcome may depend on which factor set is used.

4 Applications

While comparing factor means, correlation coefficients and absolute deviations are interesting steps of a comparison, the ultimate test is to what extent the choice of a factor set affects the result of an empirical analysis. In this section, we first analyze the performance of 41 mutual funds that focus on German stocks, then analyze how well the factors can explain the returns on 16 size/book-to-market and 16 size/momentum portfolios.

4.1 Mutual Fund Performance

We use a sample of 41 mutual funds that invest in German stocks that have been in operation from 07/1996 to 08/2011.[37] Thus, we only consider mutual funds that existed over the whole time period. This introduces a survivorship bias but satisfies the purpose of this application – to demonstrate that the usage of different data sets of Fama/French factors for Germany can make a huge difference.

We estimate the abnormal performance with respect to the four-factor model by Carhart (1997):[38]

$$R_{it} - R_t^f = \alpha_i + \beta_i\left(R_t^M - R_t^f\right) + s_i R_t^{SMB} + h_i R_t^{HML} + w_i R_t^{MOM} + \varepsilon_{it}\,, \quad (1)$$

where R_{it} is the return of fund i and R_t^f the risk free rate in month t. R_t^M denotes the return on the market portfolio. The returns on the factor portfolios are given by R_t, with the subscripts SMB, HML and MOM. The white noise error term is ε_{it} and the subscript i denotes the estimates of the abnormal performance a_i, and the factor loadings β_i, s_i, h_i and w_i for fund i.

We run this regression for the 41 funds individually as well as for an umbrella fund (fund of funds) for the whole time period of 07/1996 to 08/2011 and three sub-periods. The rate of return on the umbrella fund is the market value weighed return (market value as of December of the previous year) of the individual funds' returns. Under the assumption that the four factor model appropriately describes the cross-section of stock returns, an abnormal performance exists if a_i significantly deviates from zero.

---

[37] The mutual fund data is precisely explained in Lehmann (2013). Our sample builds on his final sample of 128 funds. Important to note, the mutual fund data is from the BVI and the rates of return of funds thus contains the corporate income tax credit (see section 2.6). To be correct, only data sets that include the corporate income tax credit should be used to evaluate the performance of these funds. However, since we only compare the different factor time series and none of the other suppliers provides a data set with the tax credit, we omit this point and only use our time series without the corporate income tax credit.

[38] The one factor version of the model goes back to Jensen (1966) and Black/Jensen/Scholes (1972) [BJS]. Fama/French (1993) proposed the three-factor model, Carhart (1997) added momentum. The assumptions, strengths and weaknesses of the test are discussed in many places, recently in Brückner et al. (2012).

[Table 7]

Table 7, on the left-hand side, shows the means of the annualized alphas and the means of the adjusted R^2 of the 41 regressions we conduct for each data set. Also shown for each factor set is the number of individual funds that have a significant negative or positive alpha (on at least a 10 % level). When looking at the full time period of 07/1996 to 08/2011, we obtain for all factor sets except two (Frazzini and Schmidt et al. with a size breakpoint of 0.5) a negative mean alpha, which is in line with our expectations. The mean adj. R^2 is close to 0.9 for four time series (our two 'ALL' series, our 'TOP' series and the Artmann et al. factor time series). All of these four can identify several funds that showed a statistically significant under- or over-performance. The adj. R^2 for the two factor series of Schmidt et al. are close to 0.8 and can at least identify one statistically significant underperformer. The adj. R^2 for Marmi/Poma is a bit below 0.5 and this factor set cannot identify a statistically significant performance. Nevertheless, the estimate for the mean alpha based on Marmi/Poma is very close to the mean alphas of several other data sets. Thus, the weaknesses of this data set can only be detected in a comparison with the other data sets. The high positive mean of alphas of Frazzini's data set (1.57, annualized) is out of line with the others and with our expectations. It may partly be related to the low mean of the market proxy of this data set.

The right-hand side of Table 7 looks at the combined value-weight portfolio of the 41 funds (Fund of Funds). The adj. R^2 are all a bit higher, but have the same patterns across the different factor time series as the mean adj. R^2 based on the individual funds, in the full period and the three sub-periods. Again, large differences between the estimated alphas occur, ranging from -1.44 (Artmann et al.) to 1.52 (Frazzini). None of the alphas are statistically significant. Although, some factor sets result in statistically significant factor loadings, especially in the sub-periods, while others do not show any significance.

The results within the three sub-periods are even more diverse than in the full period, for both, the analyses of individual funds and the analyses of the fund of funds. In the 5-year sub-period 07/1996 to 06/2001 several factor sets produce negative mean alphas, while others produce positive mean alphas. The mean alpha based on the factor series of Artmann et al. is most negative (-2.47), which we blame on the DAFOX error in 1998.[39] The mean alpha based on Frazzini is 1.95, Hanauer et al. show 1.80, that of Schmidt et al. with a size breakpoint 0.5 is 1.17 (all annualized). Artmann et al.'s factors indicate that seven individual funds had a statistically significant negative alpha, while none had a positive alpha. Other factor sets show zero funds with a negative alpha. Marmi/Poma's factors lead to a mean alpha of (plus) 9.55

[39] See Section 2.7. Note that Bessler/Drobetz/Zimmermann (2009), in their analysis of the conditional performance of German equity mutual funds note: "Again, we observe that the DAFOX index produces lower alphas than the other indexes […]".

%, which is totally out of line with the others. The four time series that had the highest adj. R^2 in the full period also have the highest adj. R^2 in this and in the other sub-periods.

In the time period 07/2001 to 06/2006 all factor sets lead to a negative alpha. All, except Marmi/Poma and Hanauer et al., are very similar. The numbers of statistically significant alphas for individual funds vary between 23 and 0. In the time period 07/2006 to 08/2011, some alphas are again negative, while others are positive.

The exploratory analysis of the 41 mutual funds that focus on German stocks clearly demonstrates that at present, a researcher must be very careful when analyzing mutual fund performance in Germany with the four-factor model. Different factor sets may result in very different estimates for the abnormal performance of individual funds and for their average abnormal performance. In the short run, the only way to reach conclusions about the performance of German mutual funds based on a three or four factor model seems to be the use of alternative factor sets in the analysis to get a feeling for the robustness of the results.

Possibly, a one-factor model using only a proxy for the market portfolio is a “better” indicator for the performance of mutual funds than the four factor models we looked at. We have four time series for the market portfolio that are very similar and probably would produce very similar conclusions. Also, the uncertainty about the quality of the SMB, HML and MOM factors of the different providers would be avoided. Finally, Brückner/Lehmann/Stehle (2013) indicate that the CAPM in its standard version seems to work well for Germany.

4.2 Size/Book-to-Market and Size/Momentum Portfolios

In our second application, we take the universe of all German non-financial stocks in our database which are listed in the top and middle segments of the FSE and the *Neuer Markt* (during its existence). That is, the stocks, on which our SMB and HML estimates in the factor sets (ALL) are based. From it, we create 4x4 size/book-to-market and 4x4 size/momentum portfolios. We first sort all firm observations from the top segment according to size, then within their size quartile according to their book-to-market or momentum characteristic (dependent sort). Using the breakpoints from the top segment, we allocate the stocks from the middle segment and the *Neuer Markt* to the portfolios. As a result, we have 16 size/book-to-market and 16 size/momentum portfolios that do not include financials. We then regress the excess return of these portfolios on the different Fama/French factor time series. Note that this procedure favors our factor set, since the test assets and the factors are based on the same database.

For each factor set, we test the null hypothesis that all 16 alphas are jointly zero by employing the Gibbons/Ross/Shanken (1989) test (GRS test). The GRS test in this context is a major improvement over the procedure suggested by BJS. It is based on the same assumptions as the BJS test procedure, stable factor loadings and homoscedastic errors plus the assumption that the variance-covariance of the residuals is stable over time. Since these assumptions may be violated to some extent, the p-values of the tests may be biased downwards, especially when the test covers a long time period.[40]

[Table 8]

Table 8 shows the results for the whole time period of 07/1996 to 12/2011 as well as for three sub-periods, separately for the 16 size/book-to-market and the 16 size/momentum portfolios. Note that we report in addition to the mean alpha the mean value of the absolute alphas, because the latter is an important input factor in the GRS test.

The nine factor sets result in very different p-values when we look at the full period for the 16 size/book-to-market portfolios. In the analysis of the full time period, all p-values (except Marmi/Poma) vary between 0.04 and 0.30. In all three sub periods the p-values are very similar, in the sub-period 07/2001 to 06/2006 all p-values are below 0.02. In the analysis of the 16 size/momentum portfolios the alternative factors sets also produce GRS p-values that are very similar. In the full time period all are essentially zero, from 07/2001 to 06/2006 all are below 0.01. For the sub-periods 07/1996 to 06/2001 they are all below 0.14 and from 07/2006 to 12/2011 they are all below 0.11, except Hanauer et al. Important to note, in both portfolio sorts, the means of the adjusted R^2s are always very close to each other.

A possible interpretation of these results is, that because of the high power of the GRS test, it does not matter much which factor set is used, all come nearly to the same conclusion. Looking at Jensen's alphas of individual portfolios separately involves test statistics with a much lower power, so the different factor sets may result in differing test results.

5 Conclusion

Germany is very fortunate with respect to the number of Fama/French factor sets that are available. As a consequence, researchers as well as practitioners can choose the factor sets that fit best to their data with respect to factor construction and the underlying data set. In addition, they can check the robustness of their results by alternatively using several factor time series.

This paper sheds light onto the question of which data sets of Fama/French factors for the German capital market have a good quality and which do not. We mainly analyze the data supplied by Artmann et al., Hanauer et al., Schmidt et al., Marmi/Poma, Frazzini and a data

[40] The assumptions, strengths and weaknesses of the GRS test are discussed in many places, e.g. Brückner et al. (2013).

set recently made available by us. The data sets differ with respect to (1) the included stock exchanges, (2) the included exchange segments, (3) whether and how they deviate from the calculation procedure suggested by Fama/French, (4) their treatment of German stock market peculiarities, (5) the underlying data bases, (6) whether the proxy for market portfolio is a publically available index or is calculated on the basis of the underlying data set and (7) by the covered time period. As a result, the Fama/French factor time series diverge considerably between the five suppliers. We conclude that the major cause for differences is the quality of the underlying stock market data bases. We also find that existing quality deficits become smaller over time. During the last ten years, the best factor data sets are very similar.

We do not recommend using the data provided by Marmi/Poma. The HML time series based on MSCI portfolios differs from the other HML time series, which is not surprising, since these portfolios are constructed in a very different way. Also, the German data set by Frazzini gives the impression that it cannot compete with the other, 'major' providers of factors for Germany.

We also do not recommend using the DAFOX as a proxy for the market portfolio in 1998. Because the DAFOX return is much higher than the other proxies in this year, we suggest switching from the DAFOX to the CDAX after the official start of the CDAX in April 1993. Before that date, the CDAX should not be used, because this index does not include dividends before 1987 and the calculation procedure for the years 1987 to 1993 is not documented.

Artmann et al. is the factor data set, which, by construction and the underlying data, is very similar to our data set 'ALL'. Its mean return on the market portfolio between 1996 and 2011 is considerably higher than ours - we believe mainly because they use the DAFOX in 1998. An additional weakness of this factor set is that they include the successful firms of the middle segment of the FSE in their calculations of SMB, HML and MOM, but not the unsuccessful ones. Consequently, the relatively small differences between the two SMB and HML data sets are difficult to evaluate. The MOM factor time series provided by Artmann et al. are somewhat out-of-line with all other MOM factor series.

The factor data provided by Schmidt et al. for the years before 1996 we also do not recommend. Before that year the combination of Datastream and Worldscope data produces a much smaller number of stocks for which the book-to-market ratio is available than in later years. In addition, for the years before 1990, Datastream data has many weaknesses, see Brückner (2013). After 1996, Schmidt et al. include considerably more stocks in their factor calculation than all other factor providers do. This seems to be related to the inclusion of a large number of stocks from the lowest segment, which we think is not beneficial. However, it may suit some, but probably not all users. Nevertheless, for most users, their series with a size breakpoint of 0.8 seems to be much more appropriate than the one with 0.5.

Hanauer et al., who also use the combination of Datastream and Worldscope data, made a wise decision to start in 1996. Unusual is their exclusion of financials from the their market portfolio, which as a consequence, has a much higher mean return than the proxies used by other providers. The proxy for the market portfolio offered by French is fully in line with the CDAX, the MSCI index for Germany, and our time series.

We recommend our time series because we offer a choice with respect to market segments and breakpoints. To users whose basic data consists mainly of stocks listed in the top segment of the FSE we recommend the factor time series ‘TOP’. To users whose basic data includes the *Neuer Markt* we recommend the factor time series ‘ALL’ with breakpoints from the ‘TOP’ data set. Another reason that speaks for our data is that we document our checks of the underlying stock database in greater detail than the other factor providers.

To the three suppliers of international time series we recommend to include a local expert for each country that is covered and to check the quality of the underlying databases.

6 Tables and Figures

Table 1: Overview of Suppliers of Fama/French Factors for the German Stock Market

Authors	Artmann/Finter/Kempf/ Koch/Theissen	Brückner/Lehmann/ Schmidt/Stehle	Frazzini[A]	Hanauer/Kaserer/Rapp	Marmi/Poma	Schmidt/Schrimpf/ von Arx/Wagner/Ziegler
Universities	Cologne	Humboldt	Stern School of Business, New York	Munich (TUM), Marburg	Scuola Normale Superiore, Pisa	Zurich, ETH Zürich
Online Data Library	http://www.cfr-cologne.de/	http://www.wiwi.hu-berlin.de/professuren/bwl/bb	http://www.econ.yale.edu/~af227/data_library.htm	http://www.fm.wi.tum.de/en/research/data	http://homepage.sns.it/marmi/Data_Library.html	http://www.bf.uzh.ch/cms/publikationen/risk-factor-database_168_1633.html
			Panel A: Coverage/Data			
Time Period	07/1962 - 12/2011	07/1958 - 06/2013	07/1990 - 09/2013[B]	08/1996 - 01/2012[C]	07/1988 - 03/2013[D]	07/1984 - 06/2012[E]
Exchanges	Frankfurt	Frankfurt	unknown	Frankfurt	Frankfurt	probably all
Markets/ Segments/ Stocks	AM, NM, GM[F]	1) 'TOP': AM 2) ‚ALL': AM, GM, NM, RM	unknown	CDAX	unknown	German Stocks in Datastream which are "major"[G]
Data Sources	KKMDB, *Saling/Hoppenstedt Aktienführer*	Several[H]	CRSP, XpressFeed Global	Datastream, Worldscope	Factset	Datastream, Worldscope
R_m	DAFOX, CDAX	From the Sample	From the Sample	From the Sample	From the Sample	From the Sample
Risk Free Rate	1-Month Money Market rate	1-Month Money Market rate, EURIBOR	Not available	FIBOR, EURIBOR	„Germany 3 months Treasury bill rate"	1-Month Frankfurt Banks (Middle Rate)
			Panel B: Ingredients			
Financials	Factors: No / R_m: Yes	Factors: No / R_m: Yes	Factors: Yes / R_m: Yes	Factors: No / R_m: **No**	Factors: No / R_m: No	Factors: Yes / R_m: Yes
Dual Class Shares	Preferred and Common Stocks	Preferred and Common Stocks	Only Common Stocks	Preferred and Common Stocks	Only Common Stocks	Only Common Stocks
Tax Credits	No	Yes/No	No	No	No	No
Penny Stocks	Included	Delete Stocks if price < 1.00 AND whose ME of firm is < 5 mln.	unknown	Included	Stocks included if avg. traded vol. > 1,000 "on the 5 prior days"	Delete stocks with prices < 1.00 of the domestic currency

Panel C: Factor Calculation/Breakpoints						
Observations involved in portfolios	Size/BM and Size/MOM portf. from the same Sub-Sample	Size/BM and Size/MOM portf. from the same Sub-Sample	unknown	Size/BM and Size/MOM portf. from the same Sub-Sample	Size/BM and Size/MOM portf. from different Sub-Samples	Size/BM and Size/MOM portf. from different Sub-Samples
Size breakpoints	Median	Median of 'TOP' or 'ALL'	0.8	Median	Median	0.8 and Median
BM breakpoints	0.3 and 0.7	0.3 and 0.7 of 'TOP' or 'ALL'	0.3 and 0.7	0.3 and 0.7	0.3 and 0.7	0.3 and 0.7
Momentum breakpoints/ calculation	0.3 and 0.7 / t-12 to t-2	0.3 and 0.7 of 'TOP' or 'ALL' / t-12 to t-2	0.3 and 0.7? / t-12 to t-2	0.3 and 0.7 / t-12 to t-2	0.3 and 0.7 / t-12 to t-2	0.3 and 0.7 / t-12 to t-2
Panel D: Provided Data						
Monthly/ Daily Data	Yes/No	Yes/Yes	Yes/No	Yes/No	Yes/No	Yes/No
Data for Germany	- Factors - Decile Portf. - 4x4 Portf.	- Factors - Decile Portf. - 4x4 Portf. - 2x3 size/BM - 2x3 size/MOM - Breakpoints	- Factors	- Factors	- Factors - 2x3 size/BM - 2x3 size/MOM	- Factors with size breakpoints 0.5 and 0.8
Multi-Country Data	No	No	Yes	No	Yes	Yes
Updates	Not Updated	Regularly Updated	Regularly Updated	unknown	Regularly Updated	Sequential Updates, No retrospective Updates

[A] The data provided by Frazzini is U.S.dollar returns. We use the exchange rates by the Deutsche Bundesbank (USD-DM, time series BBK01.WT5009, and from 1999 onwards USD-EUR, time series BBK01.WT5636) to calculate local returns. Note, that he provides a 'market factor' (minus the one-month T-Bill). The HML time series we use in our comparison is the "standard" one as proposed by Fama/French (1992, 1993). We do not discuss the additional series "HML (devil)", Asness/Frazzini (2013).

[B] The return on the market portfolio is already available from 01/1986 and the momentum factor from 01/1987.

[C] Returns are calculated on the basis of the closing prices of the first day of a month, which is unusual and may create a bias in an application as well as in our comparison with other factor providers. E.g., the row "01.01.2012" expresses the factor returns of 01 December 2011 to 01 January 2012.

[D] The momentum factor commences in 12/1988 and is surprisingly not available for a number of months in 1994, 1995 and 1996.

[E] The momentum factor is only available from 07/1987 onwards.

[F] *Geregelter Markt* stocks (GM) only, if a stock were part of the *Amtlicher Markt* (AM) or *Neuer Markt* (NM) at a later or earlier time.

[G] Domestic equity with Datastream code MAYOR="Y" and TYPE="EQ" (equity). Thus, they most likely consider stocks listed in all segments. However, this selection is based on the research lists WSCOPEBD, FGER1, FGER2, DEADBD1 and DEADBD2. With this selection, Schmidt et al. probably miss some German equities as, in addition, there are the lists FGERDOM, FGKURS and DEADBD3 to DEADBD6. Brückner (2013) discuss these selection issues in more detail.

[H] *Saling/ Hoppenstedt Aktienführer & Kurstabellen*, *Börsenzeitung*, HBDA, DFDB, KKMDB, Datastream, Worldscope

Table 2: Number of Listed vs. Number of Observations Included in the Factor Time Series

This table shows the total number of stocks/observations included in the market portfolio, size/book-to-market and size/momentum portfolios for different data suppliers as well as statistics of stock market listings in Germany (German stocks only, source: DAI Factbooks). The number of German stocks in Datastream and the CDAX are our estimates. For the number of observations included in the market portfolio, we do not have the data of Marmi/Poma available. The market portfolio of Artmann et al. is the DAFOX, for which we only have limited data about the number of stocks included (see Göppl/Schütz (1995)). Also, Artmann et al. only provide 5-year intervals for their size/book-to-market and size/momentum portfolios. Frazzini does not provide a time series of the number of included observations. "na" stands for not available.

			Stock Listings						Factor Providers																	
				At the FSE					R_m						Size/Book-to-Market						Size/Momentum					
Time	Datastream	CDAX	At least one German Exchange	Top Segment	Middle Segment	Lowest Segment	Neuer Markt	Sum	Our: TOP	Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.	French	Our: TOP	Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.	Marmi/Poma	Our: TOP	Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.	Marmi/Poma
1958-12	-	-	na	na	na	na	-	na	254	254	-	-	-	-	175	175	-	-	-	-	175	175	-	-	-	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1987-12	269	na	679	na	na	na	-	na	261	282	209	-	263	98	175	175	116	-	258	-	175	175	195	-	258	-
1988-12	623	na	706	na	na	na	-	na	267	308	406	-	268	101	187	194	142	-	258	93	187	194	211	-	258	86
1989-12	659	na	749	na	na	na	-	na	280	334	481	-	282	113	193	220	282	-	258	130	193	220	468	-	258	116
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1993-12	735	327	796	na	na	na	-	na	325	403	556	-	na	131	222	286	387	-	317	235	222	286	549	-	317	225
1994-12	746	334	810	na	na	na	-	na	334	418	563	-	na	131	221	286	440	-	317	120	221	286	550	-	317	72
1995-12	758	347	812	na	na	na	-	na	346	433	585	-	na	130	229	293	438	-	317	76	229	293	565	-	317	78
1996-12	758	355	802	na	na	na	-	na	354	434	616	203	na	126	225	290	444	203	363	93	225	290	575	203	363	91
1997-12	783	357	817	na	na	na	13	535?	356	448	636	207	na	127	230	293	567	207	363	99	230	293	612	207	363	103
1998-12	844	521	883	323	75	88	54	540	378	512	699	215	na	123	245	313	586	215	363	229	245	313	629	215	363	217
1999-12	1000	671	931	359	88	93	168	708	409	665	908	317	na	115	262	351	651	317	569	236	262	351	709	317	569	248
2000-12	1115	788	1065	365	95	160	283	903	407	781	1103	408	na	98	287	457	823	408	569	284	287	457	903	408	569	287
2001-12	1117	792	1075	359	118	163	272	912	398	787	1148	542	na	109	290	586	978	542	569	315	290	586	1095	542	569	346
2002-12	1054	755	1011	333	184	152	198	867	369	752	1040	567	na	104	277	566	870	567	524	334	277	566	1020	567	524	292
2003-12	1012	722	976	308	376	145 ┘		829	342	719	981	507	na	119	241	506	795	507	524	266	241	506	975	507	524	260
2004-12	1007	694	979	293	367	156	-	816	324	691	925	462	na	113	227	480	741	462	524	236	227	480	902	462	524	258
2005-12	1021	678	976	294	354	187	-	835	316	673	908	441	na	110	215	463	715	441	482	231	215	463	878	441	482	262
2006-12	1139	687	1103	308	348	322	-	978	329	681	962	441	na	111	216	470	720	441	482	255	216	470	864	441	482	260
2007-12	na	685	1171	658 ┘		387	-	1045	329	684	1013	437	na	89	224	469	779	437	na	293	224	469	937	437	na	284
2008-12	na	670	1178	638	-	416	-	1054	320	668	970	445	na	1013	242	483	717	445	na	265	242	483	943	445	na	272
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2013-06	na	518	na	na	-	na	-	na	242	516	-	-	-	-	186	389	-	-	-	-	186	389	-	-	-	-

Table 3: Descriptive Statistics of Fama/French Factors for Germany

The descriptive statistics are all based on the time period of 1996/07 to 12/2011 (186 month) for which the data of all suppliers is overlapping [the data set by Schmidt et al. with size breakpoint 0.8 and 0.5; the approximation by MSCI indices is described in Appendix G; Our data sets 'TOP' includes only the top segment, 'ALL' also the middle segment and the *Neuer Markt*, both without the tax credit]. BP stands for breakpoint. ***/**/* show significance on a 1/5/10 % level.

	Min	0.25	Median	0.75	Max	Mean	Sd	Kurt	Skew
					R_m- R_f				
Our: TOP	-21.60	-2.87	1.15	4.17	17.44	0.46	5.85	4.35	-0.52
Our: ALL	-21.59	-2.99	1.35	4.10	17.32	0.42	5.91	4.21	-0.49
Schmidt et al.	-16.46	-2.75	0.53	4.08	14.70	0.39	5.61	3.51	-0.56
Hanauer et al.	-15.49	-2.61	1.25	4.60	13.83	0.56	5.92	3.17	-0.53
Artmann et al.	-20.98	-2.92	1.31	4.35	17.37	0.53	5.99	4.44	-0.53
Marmi/Poma	-11.45	-1.72	0.63	2.55	22.02	0.31	4.11	7.31	0.54
French	-24.17	-2.94	0.75	4.54	19.12	0.44	6.23	4.49	-0.53
MSCI	-25.20	-3.17	1.05	4.93	20.73	0.45	6.74	4.47	-0.51
CDAX	-24.12	-3.24	1.11	4.38	19.80	0.40	6.42	4.53	-0.56
Frazzini	-17.06	-2.65	0.71	3.65	15.97	0.28	5.67	3.73	-0.39
					SMB				
Our: TOP	-14.02	-3.19	-0.48	2.67	9.61	-0.32	4.06	3.06	-0.18
Our: ALL [BPs: TOP]	-14.55	-3.39	-0.48	2.31	8.58	-0.48*	3.84	3.10	-0.21
Our: ALL	-15.03	-3.45	-1.31	1.84	11.33	-0.80***	4.08	3.38	0.02
Schmidt et al. [size-BP: 0.5]	-10.50	-3.02	-0.26	1.98	12.98	-0.28	3.90	3.41	0.23
Schmidt et al. [size-BP: 0.8]	-10.26	-2.43	-0.30	1.76	9.79	-0.27	3.39	3.29	0.06
Hanauer et al.	-14.30	-3.36	-0.65	2.41	10.36	-0.71**	4.14	3.03	-0.23
Artmann et al.	-12.96	-3.05	-0.46	2.25	10.61	-0.55*	4.00	3.36	-0.21
Marmi/Poma	-11.74	-2.76	-0.62	1.36	24.91	-0.56*	3.99	11.49	1.43
MSCI	-14.33	-2.01	-0.06	2.22	7.43	-0.09	3.57	4.77	-0.71
Frazzini	-13.66	-2.50	-0.40	2.78	12.30	-0.03	4.05	3.40	0.06
					HML				
Our: TOP	-13.16	-1.03	0.47	2.71	15.62	0.42*	3.46	6.05	-0.05
Our: ALL [BPs: TOP]	-15.95	-1.07	0.55	2.57	16.14	0.64**	3.70	7.37	0.07
Our: ALL	-12.85	-1.00	0.62	2.40	14.34	0.76***	3.51	6.02	0.34
Schmidt et al. [size-BP: 0.5]	-8.76	-1.48	0.57	2.65	11.03	0.55**	3.68	3.56	0.31
Schmidt et al. [size-BP: 0.8]	-14.74	-1.05	0.81	2.59	13.47	0.77***	3.86	5.57	-0.08
Hanauer et al.	-10.36	-1.48	0.37	2.89	11.91	0.74***	3.49	3.79	0.32
Artmann et al.	-12.24	-1.21	0.75	2.77	19.23	0.85***	3.93	6.67	0.43
Marmi/Poma	-15.15	-1.20	1.10	2.64	19.00	1.13***	4.06	6.89	0.61
MSCI	-20.75	-1.80	0.42	2.16	12.98	0.24	4.50	6.52	-0.62
Frazzini	-12.38	-2.61	0.19	3.74	15.59	0.67*	4.65	3.61	0.40
					MOM				
Our: TOP	-20.40	-1.29	1.04	3.54	21.81	1.32***	5.26	6.94	-0.21
Our: ALL [BPs: TOP]	-21.29	-0.80	1.25	3.61	20.30	1.34***	5.66	6.22	-0.30
Our: ALL	-32.93	-1.32	1.34	3.73	22.43	1.05**	6.63	8.74	-1.11
Schmidt et al. [size-BP: 0.5]	-32.91	-0.80	1.68	3.63	18.97	1.31***	6.63	9.24	-1.31
Schmidt et al. [size-BP: 0.8]	-37.54	-1.18	1.39	4.33	21.95	1.27**	7.23	8.99	-1.14
Hanauer et al.	-40.71	-1.64	1.53	3.86	19.99	1.19**	7.02	11.09	-1.48
Artmann et al.	-21.51	-0.10	1.52	3.30	17.55	1.44***	5.21	7.32	-0.68
Marmi/Poma	-23.19	-1.01	1.28	4.23	20.68	1.60***	6.12	5.61	-0.10
Frazzini	-23.71	-2.05	1.52	4.34	30.62	1.24**	7.25	5.95	-0.12

Table 4: Time Series Analyses of the Fama/French Factors for Germany

This table shows R²s and test statistics (p. values) for two-sample (pair-wise) mean and mean rank tests between single time series of different Fama/French data sets [the data set by Schmidt et al. with size breakpoint 0.8; the approximation by MSCI indices is described in Appendix G; Our data set 'ALL', without tax credit and breakpoints from the top segment]. The R² is from an OLS regression where the left hand variable is a time series of column one, explained by a competing time series within that row. The analysis is based on the overlapping time frame of 07/1996 to 12/2011 (186 month). Values below 0.5 (0.3) for the panel of R²s and values below 0.1 (0.05) for the two panels of p-values of tests are colored (dark) grey.

	R²									T-test (pair wise) [p. value]									Wilcoxon Signed-Rank Test [p-value]								
	Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.	Marmi/ Poma	French	MSCI	CDAX	Frazzini	Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.	Marmi/ Poma	French	MSCI	CDAX	Frazzini	Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.	Marmi/ Poma	French	MSCI	CDAX	Frazzini
R_m - R_f																											
Our: ALL	1.00	0.89	0.85	0.97	0.50	0.97	0.97	0.98	0.91		0.86	0.43	0.17	0.71	0.83	0.77	0.74	0.28		0.76	0.43	0.15	0.70	0.90	0.94	0.97	0.24
Schmidt et al.	0.89	1.00	0.96	0.86	0.48	0.88	0.88	0.89	0.82	0.86		0.08	0.43	0.77	0.79	0.76	1.00	0.52	0.76		0.06	0.34	0.84	0.96	0.97	0.93	0.32
Hanauer et al.	0.85	0.96	1.00	0.82	0.42	0.85	0.84	0.85	0.80	0.43	0.08		0.87	0.45	0.51	0.60	0.38	0.16	0.43	0.06		0.75	0.51	0.33	0.34	0.24	0.07
Artmann et al.	0.97	0.86	0.82	1.00	0.49	0.94	0.96	0.97	0.88	0.17	0.43	0.87		0.49	0.42	0.49	0.13	0.11	0.15	0.34	0.75		0.44	0.44	0.51	0.61	0.09
Marmi/Poma	0.50	0.48	0.42	0.49	1.00	0.47	0.46	0.49	0.45	0.71	0.77	0.45	0.49		0.70	0.70	0.80	0.93	0.70	0.84	0.51	0.44		0.78	0.68	0.80	0.83
French	0.97	0.88	0.85	0.94	0.47	1.00	0.99	0.98	0.90	0.83	0.79	0.51	0.42	0.70		0.84	0.54	0.28	0.90	0.96	0.33	0.44	0.78		0.54	0.78	0.20
MSCI	0.97	0.88	0.84	0.96	0.46	0.99	1.00	0.99	0.90	0.77	0.76	0.60	0.49	0.70	0.84		0.32	0.31	0.94	0.97	0.34	0.51	0.68	0.54		0.52	0.17
CDAX	0.98	0.89	0.85	0.97	0.49	0.98	0.99	1.00	0.91	0.74	1.00	0.38	0.13	0.80	0.54	0.32		0.43	0.97	0.93	0.24	0.61	0.80	0.78	0.52		0.29
SMB																											
Our: ALL	1.00	0.56	0.65	0.74	0.04		0.41		0.33		0.28	0.21	0.62	0.82		0.09		0.10		0.41	0.17	0.82	0.68		0.07		0.07
Schmidt et al.	0.56	1.00	0.56	0.51	0.04		0.33		0.30	0.28		0.03	0.17	0.40		0.44		0.38	0.41		0.06	0.36	0.31		0.18		0.53
Hanauer et al.	0.65	0.56	1.00	0.77	0.07		0.24		0.26	0.21	0.03		0.30	0.68		0.03		0.02	0.17	0.06		0.47	0.88		0.03		0.03
Artmann et al.	0.74	0.51	0.77	1.00	0.07		0.33		0.33	0.62	0.17	0.30		0.99		0.07		0.06	0.82	0.36	0.47		0.64		0.05		0.02
Marmi/Poma	0.04	0.04	0.07	0.07	1.00		0.02		0.01	0.82	0.40	0.68	0.99			0.19		0.19	0.68	0.31	0.88	0.64			0.11		0.08
MSCI	0.41	0.33	0.24	0.33	0.02		1.00		0.16	0.09	0.44	0.03	0.07	0.19				0.86	0.07	0.18	0.03	0.05	0.11				0.81
HML																											
Our: ALL	1.00	0.44	0.44	0.47	0.15		0.20		0.29		0.56	0.66	0.35	0.12	0.56	0.21		0.93		0.64	0.91	0.41	0.42	0.72	0.09		0.66
Schmidt et al.	0.44	1.00	0.37	0.42	0.07		0.16		0.20	0.56		0.87	0.76	0.32	0.85	0.12		0.75	0.64		0.69	0.78	0.49	0.85	0.01		0.31
Hanauer et al.	0.44	0.37	1.00	0.49	0.06		0.12		0.21	0.66	0.87		0.60	0.25	0.78	0.15		0.83	0.91	0.69		0.53	0.40	0.67	0.18		0.50
Artmann et al.	0.47	0.42	0.49	1.00	0.18		0.16		0.22	0.35	0.76	0.60		0.37	0.96	0.08		0.58	0.41	0.78	0.53		0.75	0.85	0.20		0.29
Marmi/Poma	0.15	0.07	0.06	0.18	1.00		0.04		0.04	0.12	0.32	0.25	0.37		0.48	0.03		0.25	0.42	0.49	0.40	0.75		0.92	0.04		0.19
MSCI	0.20	0.16	0.12	0.16	0.04		1.00		0.10	0.21	0.12	0.15	0.08	0.03	0.07			0.28	0.09	0.01	0.18	0.20	0.04	0.05			0.75
MOM																											
Our: ALL	1.00	0.71	0.74	0.63	0.33				0.67		0.80	0.56	0.68	0.52				0.75		0.58	0.95	0.66	0.40				0.58
Schmidt et al.	0.71	1.00	0.83	0.66	0.34				0.61	0.80		0.71	0.57	0.47				0.94	0.58		0.74	0.87	0.80				0.54
Hanauer et al.	0.74	0.83	1.00	0.69	0.37				0.57	0.56	0.71		0.37	0.34				0.88	0.95	0.74		0.73	0.48				0.51
Artmann et al.	0.63	0.66	0.69	1.00	0.55				0.60	0.68	0.57	0.37		0.62				0.55	0.66	0.87	0.73		0.87				0.22
Marmi/Poma	0.33	0.34	0.37	0.55	1.00				0.32	0.52	0.47	0.34	0.62					0.44	0.40	0.80	0.48	0.87					0.34

Table 5: Significant Differences between the Fama/French Factor Time Series

This table shows the number of significant differences between the Fama/French factor time series [the data set by Schmidt et al. with size breakpoint 0.8; the approximation by MSCI indices is described in Appendix G; Our data set ‘ALL’, without tax credit and breakpoints from the top segment]. The analysis is based on the overlapping time frame of 1996/07 to 12/2011 (N=186). Within this time period, each time series is normalized so that they have mean zero and standard deviation one. A significant difference between two time series is an absolute difference of > 1 (one standard deviation). The last four columns show the number of overlapping data points for which we record significant differences. E.g. our R_m-R_f time series has two data points (month) where our time series deviates significantly (by more than one standard deviation) from two other suppliers.

	Mean number of sign. Differences	The significant differences are recorded towards				Overlapping in *X* month(s)			
		Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.	1	2	3	4
R_m - R_f									
Our: ALL	3.33	0	3	7	0	6	2	0	0
Schmidt et al.	2.33	3	0	0	4	3	2	0	0
Hanauer et al.	5.33	7	0	0	9	6	5	0	0
Artmann et al.	4.33	0	4	9	0	5	4	0	0
Marmi/Poma	*33.50*	*30*	*33*	*37*	*34*	*16*	*10*	*6*	*20*
French	*2.25*	*1*	*3*	*5*	*0*	*7*	*1*	*0*	*0*
MSCI	*3.00*	*1*	*5*	*6*	*0*	*6*	*3*	*0*	*0*
CDAX	*2.00*	*0*	*3*	*5*	*0*	*4*	*2*	*0*	*0*
Frazzini	*4.00*	*0*	*7*	*7*	*2*	*6*	*5*	*0*	*0*
SMB									
Our: ALL	17.67	0	26	16	11	25	11	2	0
Schmidt et al.	25.33	26	0	22	28	29	7	11	0
Hanauer et al.	16.33	16	22	0	11	30	8	1	0
Artmann et al.	16.67	11	28	11	0	33	7	1	0
Marmi/Poma	*64.00*	*72*	*58*	*64*	*62*	*33*	*24*	*21*	*28*
MSCI	*47.75*	*40*	*44*	*57*	*50*	*32*	*24*	*17*	*15*
Frazzini	*51.75*	*49*	*51*	*56*	*51*	*36*	*19*	*19*	*19*
HML									
Our: ALL	37.33	0	41	36	35	44	25	6	0
Schmidt et al.	41.33	41	0	47	36	44	22	12	0
Hanauer et al.	37.00	36	47	0	28	49	19	8	0
Artmann et al.	33.00	35	36	28	0	49	22	2	0
Marmi/Poma	*62.00*	*62*	*60*	*72*	*54*	*38*	*37*	*16*	*22*
MSCI	*50.00*	*49*	*47*	*57*	*47*	*41*	*17*	*19*	*17*
Frazzini	*58.50*	*59*	*50*	*64*	*61*	*46*	*27*	*22*	*17*
MOM									
Our: ALL	12.67	0	12	7	19	14	6	4	0
Schmidt et al.	11.00	12	0	5	16	17	5	2	0
Hanauer et al.	8.33	7	5	0	13	16	3	1	0
Artmann et al.	16.00	19	16	13	0	20	5	6	0
Marmi/Poma	*35.75*	*46*	*35*	*38*	*24*	*26*	*16*	*11*	*13*
Frazzini	*20.00*	*17*	*20*	*24*	*19*	*21*	*11*	*7*	*4*

Table 6: Months with Large Differences between the Fama/French Factors for Germany

This table shows selected months of Table 5 for which we find the largest spreads (maximum - minimum) between the data of the four 'major' data providers [the data set by Schmidt et al. with size breakpoint 0.8; Our data set 'ALL', without tax credit and breakpoints from the top segment]. We also provide the estimates by other times series [the approximation by MSCI indices is described in Appendix G].

	'Major' Data Providers				Maximum Spread	*Other Time Series*				
Time	Our: ALL	Schmidt et al.	Hanauer et al.	Artmann et al.		*Marmi/Poma*	*French*	*MSCI*	*CDAX*	*Frazzini*
					R_m - R_f					
1998-10	2.09	9.52	8.24	1.73	7.79	*-0.20*	*2.73*	*2.98*	*3.48*	*1.78*
1998-11	8.91	0.17	-0.32	10.60	10.92	*0.59*	*6.50*	*7.11*	*6.34*	*7.15*
1998-12	-1.77	3.86	5.29	-0.69	7.06	*-1.77*	*0.71*	*0.71*	*-0.33*	*0.16*
1999-11	4.92	6.58	11.23	4.35	6.88	*-1.65*	*5.75*	*7.07*	*6.17*	*5.41*
1999-12	15.65	8.64	11.68	14.95	7.01	*1.68*	*16.50*	*17.40*	*13.81*	*15.24*
2009-08	0.03	-4.16	-5.30	3.46	8.76	*2.67*	*0.04*	*2.88*	*3.46*	*2.30*
					SMB					
1998-10	-1.92	-6.31	-9.77	-5.13	7.85	*-2.25*		*-1.36*		*-4.26*
1998-11	-7.28	2.34	4.41	-0.15	11.69	*1.54*		*-3.15*		*-0.51*
1999-02	3.74	7.68	3.73	-0.60	8.28	*-1.85*		*4.40*		*5.91*
1999-03	1.32	-1.90	-6.44	-1.45	7.77	*-3.59*		*-1.00*		*4.42*
1999-04	-8.24	0.13	-5.01	-7.44	8.37	*9.12*		*-4.82*		*-2.28*
1999-11	-8.75	-6.13	-14.30	-12.96	8.17	*-2.75*		*-2.03*		*-8.28*
2002-07	2.83	9.79	2.27	1.88	7.91	*-8.54*		*1.49*		*5.23*
2003-01	-0.42	4.35	-5.79	-4.31	10.14	*1.02*		*0.16*		*0.45*
2003-04	-1.40	-8.00	-2.05	0.52	8.52	*-8.70*		*-5.28*		*-2.71*
					HML					
1998-11	2.40	-0.53	-3.14	-6.96	9.36	*2.19*		*1.98*		*4.03*
1999-01	-0.15	-11.04	-3.73	-11.20	11.05	*-1.27*		*10.42*		*4.64*
1999-03	-0.63	2.42	8.56	9.65	10.28	*0.55*		*6.40*		*5.54*
2000-02	-15.95	-10.19	-4.49	-11.15	11.46	*-8.11*		*-3.85*		*-9.49*
2000-11	14.54	13.36	9.33	19.23	9.90	*10.70*		*7.49*		*11.27*
2003-04	-1.87	6.97	7.58	-1.67	9.44	*-6.01*		*1.62*		*0.91*
2003-10	4.28	13.47	3.80	1.93	11.54	*0.14*		*1.18*		*2.56*
2006-04	1.37	-1.06	-8.05	0.66	9.42	*3.60*		*0.98*		*-3.16*
2009-07	1.51	1.96	6.94	-2.47	9.41	*1.03*		*0.79*		*3.97*
					MOM					
2000-01	7.80	8.52	6.97	-3.88	12.40	*-8.72*				*7.63*
2000-02	14.72	8.72	10.92	1.83	12.89	*3.55*				*7.04*
2001-06	-0.43	8.46	1.65	14.03	14.46	*18.76*				*11.70*
2001-09	8.54	19.64	10.03	5.41	14.23	*9.41*				*13.23*
2001-10	-10.35	-37.54	-40.71	-15.36	30.36	*-6.12*				*-14.77*
2001-12	5.64	6.15	14.22	16.79	11.15	*16.69*				*7.33*
2002-02	8.78	-1.96	10.60	9.23	12.55	*15.13*				*10.42*
2004-01	1.65	8.50	-3.29	5.88	11.79	*8.03*				*5.25*

Table 7: Mean coefficients of the Carhart four-factor model for German Mutual Funds

This table shows mean values of coefficients of the Carhardt (1997) four-factor model, $R_{it} - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i R_t^{SMB} + h_i R_t^{HML} + w_i R_t^{MOM} + \varepsilon_{it}$, applied to 41 mutual funds that invest in German stocks as well as for a fund of funds. The results are shown for different data sets for the whole, overlapping time period of 07/1996 to 08/2011 (182 month) and three sub-periods [the data set by Schmidt et al. with size breakpoint 0.8 and 0.5; Our data sets 'TOP' includes only the top segment, 'ALL' also the middle segment and the *Neuer Markt*, both without the tax credit]. "+" ("-") indicates the number of funds which have a positive (negative) and statistically significant alpha on at least a 10 % level. Alphas are annualized (multiplied by 12) and given in percentage. ***/**/* show significance on 1/5/10 % level. BP stands for breakpoint. Test statistics are based on Newey-West standard errors.

	Single Funds			Fund of Funds					
	Mean of Alphas	+ / 0 / -	Mean of Adj. R²	α	β	s (SMB)	h (HML)	w (MOM)	Adj. R²
	07/1996 - 08/2011								
Artmann et al.	-1.02	1 / 32 / 8	0.896	-1.44	1.06**	0.01	0.04	-0.07**	0.956
Hanauer et al.	-1.07	0 / 37 / 4	0.764	-1.38	1.02	0.19**	0.20***	-0.09*	0.804
Marmi/Poma	-0.13	0 / 41 / 0	0.476	-0.59	1.10	0.20	0.25*	-0.05	0.496
Our: TOP	-0.13	1 / 38 / 2	0.888	0.01	1.11**	0.07	0.01	-0.10**	0.942
Our: ALL [BPs: TOP]	-0.11	2 / 36 / 3	0.894	-0.12	1.07**	0.02	0.11***	-0.10***	0.946
Our: ALL	-0.80	2 / 34 / 5	0.892	-1.19	1.07**	0.01	0.15***	-0.09**	0.947
Schmidt et al. [size-BP:0.5]	0.33	0 / 40 / 1	0.790	0.31	1.07	0.11	0.11*	-0.08*	0.832
Schmidt et al. [size-BP:0.8]	-0.18	0 / 40 / 1	0.793	-0.35	1.05	0.05	0.14**	-0.07**	0.833
Frazzini	1.57	5 / 35 / 1	0.874	1.52	1.03	-0.09*	0.21***	-0.15***	0.919
	07/1996 - 06/2001								
Artmann et al.	-2.47	0 / 34 / 7	0.870	-2.91	1.00	-0.16***	0.06**	-0.09	0.948
Hanauer et al.	1.80	0 / 41 / 0	0.690	1.99	1.05	0.20	0.22***	-0.04	0.721
Marmi/Poma	9.55	0 / 41 / 0	0.361	9.49	1.15	0.36*	0.18	0.15	0.368
Our: TOP	-1.49	0 / 37 / 4	0.900	-0.86	1.04	0.02	0.07*	0.08**	0.969
Our: ALL [BPs: TOP]	0.58	0 / 40 / 1	0.906	0.64	1.00	-0.03	0.12***	0.03	0.971
Our: ALL	-0.37	0 / 37 / 4	0.903	-0.72	1.00	-0.03	0.13***	0.03	0.974
Schmidt et al. [size-BP:0.5]	1.17	1 / 40 / 0	0.752	1.53	1.10*	0.09	0.18***	-0.02	0.794
Schmidt et al. [size-BP:0.8]	-0.47	0 / 39 / 2	0.781	-0.85	1.13**	-0.02	0.28***	-0.03	0.824
Frazzini	1.95	3 / 38 / 0	0.869	1.54	0.97	-0.04	0.19***	-0.06*	0.930
	07/2001 - 06/2006								
Artmann et al.	-1.78	1 / 30 / 10	0.924	-2.60	1.09**	0.01	0.04	-0.07**	0.959
Hanauer et al.	-4.13	0 / 26 / 15	0.861	-5.24	1.11	0.09	0.09	-0.01	0.884
Marmi/Poma	-5.87	0 / 41 / 0	0.510	-6.97	0.98	0.01	0.34	-0.19	0.513
Our: TOP	-1.71	0 / 35 / 6	0.919	-2.56	1.11**	0.06	-0.05	-0.04	0.956
Our: ALL [BPs: TOP]	-1.80	0 / 31 / 10	0.926	-2.70	1.07*	0.01	0.06	-0.05	0.958
Our: ALL	-1.97	4 / 14 / 23	0.926	-3.06*	1.07	-0.01	0.08	-0.05*	0.959
Schmidt et al. [size-BP:0.5]	-1.80	1 / 38 / 2	0.884	-2.51	1.15*	0.00	0.00	-0.04	0.909
Schmidt et al. [size-BP:0.8]	-1.48	0 / 41 / 0	0.886	-1.87	1.10	-0.10	0.00	-0.04	0.911
Frazzini	-0.90	3 / 36 / 2	0.899	-1.94	1.05	-0.17**	0.30***	-0.11**	0.931
	07/2006 - 08/2011								
Artmann et al.	-1.01	1 / 34 / 6	0.924	0.05	1.09**	0.19*	0.03	-0.04	0.967
Hanauer et al.	-1.66	0 / 39 / 2	0.778	-1.09	0.99	0.41***	0.36*	-0.26***	0.812
Marmi/Poma	-1.06	0 / 40 / 1	0.622	-0.70	1.37**	0.67***	0.17	-0.10	0.642
Our: TOP	3.03	2 / 39 / 0	0.894	3.96	1.10	0.09	-0.08	-0.24***	0.934
Our: ALL [BPs: TOP]	1.83	0 / 40 / 1	0.897	2.97	1.10	0.08	-0.03	-0.23***	0.936
Our: ALL	1.10	2 / 38 / 1	0.896	2.04	1.08	0.07	0.07	-0.23***	0.937
Schmidt et al. [size-BP:0.5]	1.29	0 / 40 / 1	0.757	2.42	1.08	0.31	0.04	-0.19*	0.791
Schmidt et al. [size-BP:0.8]	1.11	2 / 38 / 1	0.761	2.27	1.10	0.43**	0.00	-0.14	0.792
Frazzini	2.70	7 / 34 / 0	0.885	3.84	1.12*	-0.03	0.11	-0.18***	0.913

Table 8: Alphas of the Carhart four-factor model for size/BM and size/momentum portfolios

This table shows mean values of alphas and adjusted R² of the Carhardt (1997) four-factor model, $R_{it} - R_t^f = \alpha_i + \beta_i(R_t^M - R_t^f) + s_i R_t^{SMB} + h_i R_t^{HML} + w_i R_t^{MOM} + \varepsilon_{it}$, applied to 16 size/BM and 16 size/momentum portfolios and the corresponding GRS test statistics. We first sort all non-financial firm observations from the top segment according to size, then within each size quartile by book-to-market or momentum (depended sort) and allocate the stocks from the middle segment and the *Neuer Markt* to the portfolios. The results are shown for different data sets for the time period of 07/1996 to 12/2011 (186 month) and three sub-periods [the data set by Schmidt et al. with size breakpoint 0.8 and 0.5; Our data sets 'TOP' includes only the top segment, 'ALL' also the middle segment and the *Neuer Markt*, both without the tax credit]. Alpha's are on a monthly basis, given in percentage. ***/**/* show significance for the GRS-Test on a 1/5/10 % level. BP stands for breakpoint.

	16 Size/Book-to-Market Portfolios				16 Size/Momentum Portfolios			
	Mean of |Alphas|	Mean of Alphas	Mean of Adj. R²	GRS-Test (p-values)	Mean of |Alphas|	Mean of Alphas	Mean of Adj. R²	GRS-Test (p-values)
				07/1996 - 12/2011				
Our: TOP	0.305**	0.163	0.685	0.040	0.494***	0.128	0.696	0.000
Our: ALL [BPs: TOP]	0.266**	0.193	0.730	0.048	0.455***	0.147	0.754	0.000
Our: ALL	0.312	0.282	0.697	0.156	0.464***	0.204	0.730	0.000
Schmidt et al. [size-BP:.5]	0.310*	0.111	0.652	0.086	0.461***	0.041	0.678	0.000
Schmidt et al. [size-BP:.8]	0.276	0.071	0.668	0.298	0.432***	-0.004	0.691	0.000
Hanauer et al.	0.265	0.174	0.669	0.293	0.445***	0.099	0.697	0.000
Artmann et al.	0.271	0.231	0.688	0.240	0.368***	0.187	0.702	0.000
Marmi/Poma	0.263	0.134	0.515	0.527	0.468***	0.043	0.526	0.000
Frazzini	0.405**	0.400	0.617	0.039	0.565***	0.340	0.644	0.000
				07/1996 - 06/2001				
Our: TOP	0.472	-0.186	0.522	0.562	0.708**	-0.255	0.592	0.022
Our: ALL [BPs: TOP]	0.267	0.041	0.557	0.302	0.604**	-0.019	0.628	0.031
Our: ALL	0.258	0.085	0.538	0.492	0.512*	0.019	0.614	0.067
Schmidt et al. [size-BP:.5]	0.314	-0.010	0.481	0.923	0.595*	-0.072	0.528	0.085
Schmidt et al. [size-BP:.8]	0.202	0.075	0.506	0.951	0.512	0.027	0.550	0.141
Hanauer et al.	0.284	0.131	0.545	0.910	0.552**	0.076	0.593	0.048
Artmann et al.	0.232	0.062	0.509	0.975	0.506**	0.035	0.552	0.026
Marmi/Poma	0.381	0.276	0.478	0.939	0.583*	0.152	0.482	0.055
Frazzini	0.259	0.063	0.456	0.948	0.487*	-0.013	0.503	0.093
				07/2001 - 06/2006				
Our: TOP	0.462**	0.256	0.722	0.017	0.556***	0.237	0.717	0.003
Our: ALL [BPs: TOP]	0.430**	0.346	0.806	0.014	0.528***	0.315	0.815	0.004
Our: ALL	0.578***	0.554	0.782	0.001	0.741***	0.502	0.801	0.000
Schmidt et al. [size-BP:.5]	0.584***	0.243	0.759	0.000	0.755***	0.222	0.789	0.000
Schmidt et al. [size-BP:.8]	0.604***	0.020	0.754	0.003	0.742***	0.015	0.783	0.000
Hanauer et al.	0.505***	0.362	0.764	0.005	0.740***	0.320	0.786	0.000
Artmann et al.	0.499***	0.420	0.776	0.000	0.607***	0.379	0.780	0.000
Marmi/Poma	0.449	0.223	0.507	0.120	0.715***	0.110	0.533	0.006
Frazzini	0.615***	0.601	0.689	0.003	0.786***	0.557	0.733	0.000
				07/2006 - 12/2011				
Our: TOP	0.412	0.386	0.793	0.142	0.618***	0.427	0.798	0.003
Our: ALL [BPs: TOP]	0.228	0.144	0.810	0.737	0.498**	0.221	0.811	0.013
Our: ALL	0.301	0.237	0.798	0.938	0.521**	0.316	0.797	0.040
Schmidt et al. [size-BP:.5]	0.286	0.134	0.709	0.948	0.487	0.142	0.709	0.111
Schmidt et al. [size-BP:.8]	0.273	0.172	0.736	0.895	0.488*	0.182	0.737	0.089
Hanauer et al.	0.270	0.028	0.753	0.927	0.375	0.092	0.743	0.312
Artmann et al.	0.288	0.024	0.784	0.411	0.407**	0.097	0.781	0.047
Marmi/Poma	0.278	0.104	0.637	0.867	0.525**	0.095	0.633	0.036
Frazzini	0.503	0.360	0.741	0.603	0.598***	0.336	0.731	0.009

Figure 1: Plot of rolling Means of the Market Excess Returns

This plot shows rolling means of the different market excess returns. Each observation is based on the prior 60 months, starting in 06/2001 (including the time period of 07/1996 to 06/2001).

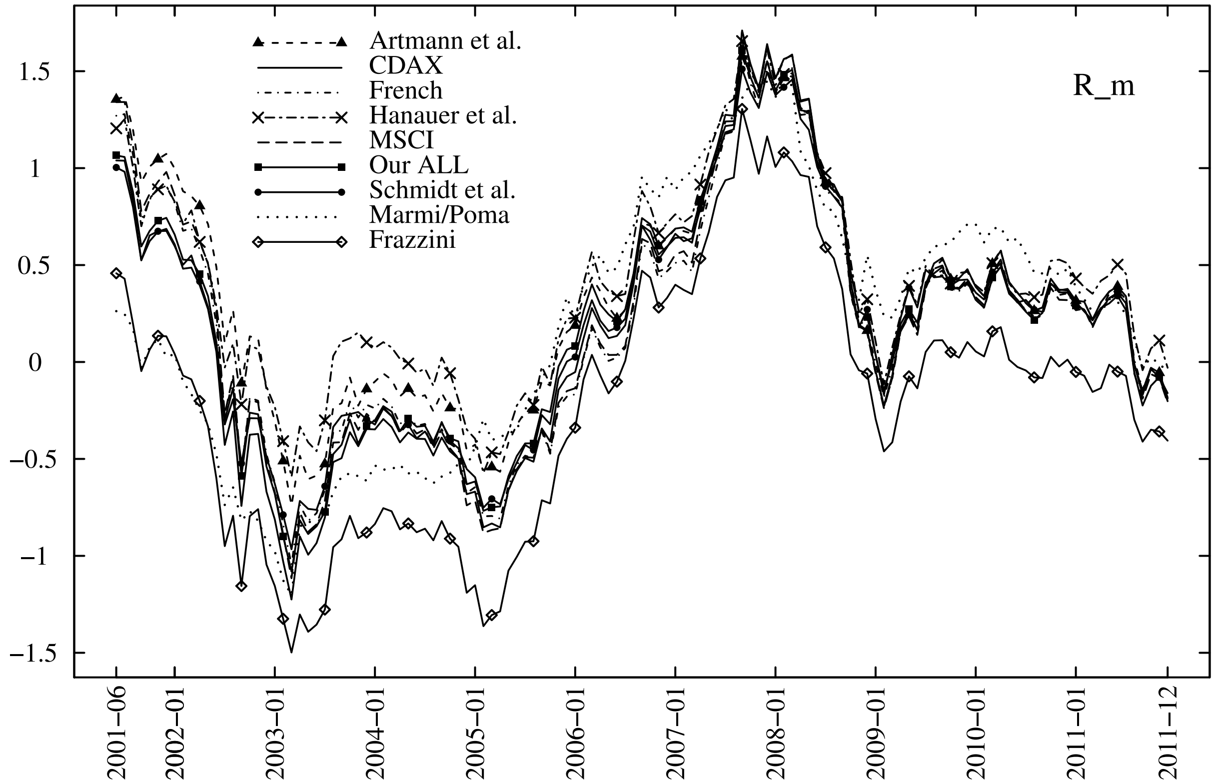


Figure 2: Plot of the SMB factor

This plot shows rolling means of the different SMB factors. Each observation is based on the prior 60 months, starting in 06/2001 (including the time period of 07/1996 to 06/2001).

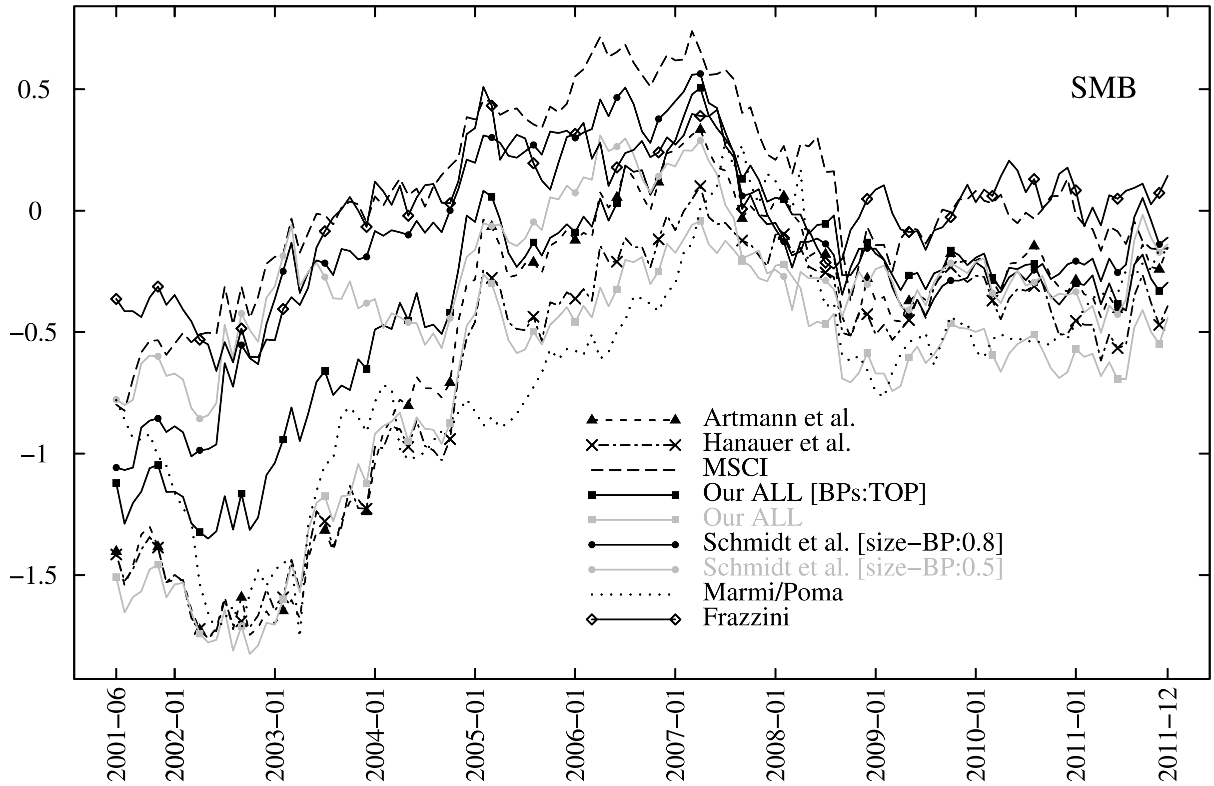


Figure 3: Plot of the HML factor

This plot shows rolling means of the different HML factors. Each observation is based on the prior 60 months, starting in 06/2001 (including the time period of 07/1996 to 06/2001).

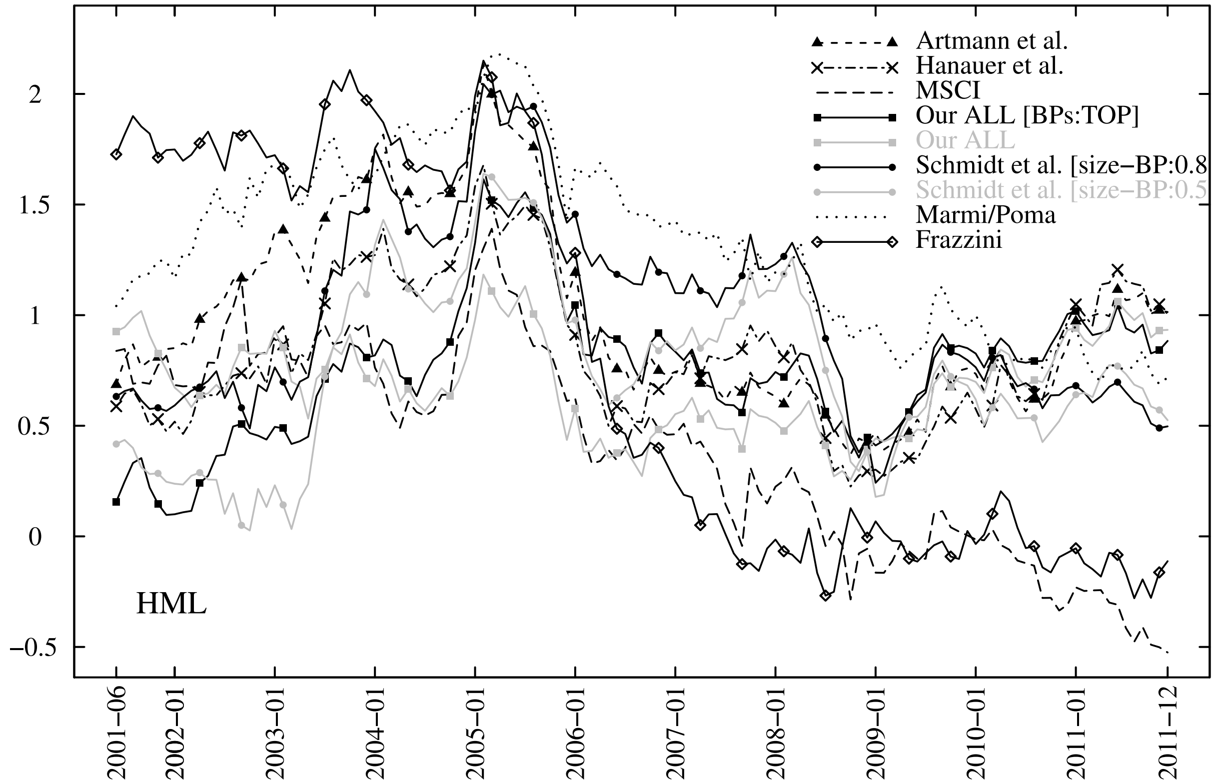
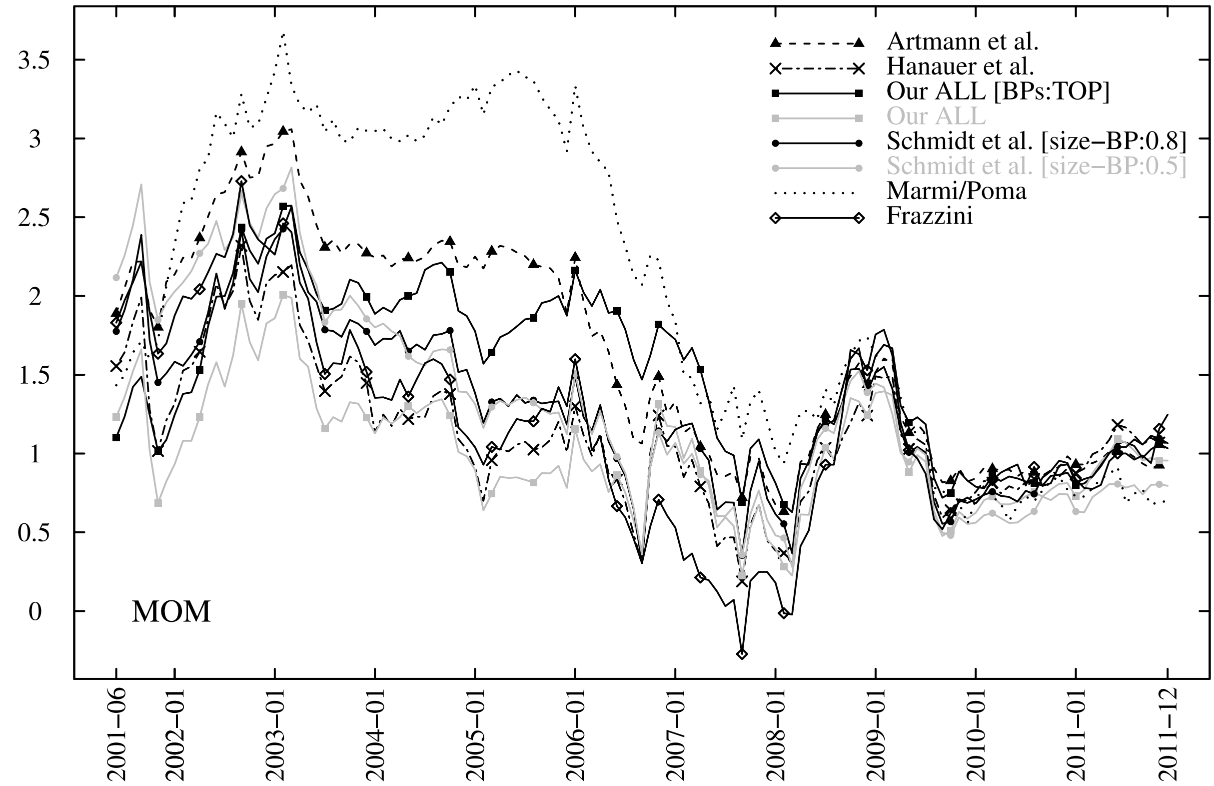


Figure 4: Plot of the MOM factor

This plot shows rolling means of the different MOM factors. Each observation is based on the prior 60 months, starting in 06/2001 (including the time period of 07/1996 to 06/2001).



7 References

Amel-Zadeh, A. (2011): "The Return of the Size Anomaly: Evidence from the German Stock Market", *European Financial Management*, 17(1), 145-182.

Ang, A./Hodrick, R. J./Xing, Y./Zhang, X. (2009): "High idiosyncratic volatility and low returns: International and further U.S. evidence", *Journal of Financial Economics*, 91(1), 1-23.

Annaert, J./De Ceuster, M./Verstegen, K. (2013): "Are extreme returns priced in the stock market? European evidence", *Journal of Banking & Finance*, 37(9), 3401-3411,

Artmann, S./Finter, P./Kempf, A. (2012a): "Determinants of Expected Stock Returns: Large Sample Evidence from the German Market", *Journal of Business Finance & Accounting*, 39(5-6), 758-784.

Artmann, S./Finter, P./Kempf, A./Koch, S./Theissen, E. (2012b): "The Cross-Section of German Stock Returns: New Data and New Evidence", *Schmalenbach Business Review*, 64, 20-43.

Asness, C. S./Moskowitz, T. J./Pedersen, L. H. (2013), "Value and Momentum Everywhere", *The Journal of Finance*, 68, 929-985.

Asness, C./Frazzini, A. (2013): "The Devil in HML's Details", *Journal of Portfolio Management*, 39(4), 49-68.

Barber, B. M./Lyon, J. D. (1997): "Firm size, book-to-market ratio, and security returns: A holdout sample of financial firms", *Journal of Finance*, 52, 875-883.

Bessler, W./Drobetz, W./Zimmermann, H. (2009): "Conditional performance evaluation for German equity mutual funds", *The European Journal of Finance*, 15(3), 287-316.

Black, F./Jensen, M. C./Scholes, M. (1972): "The capital asset pricing model: Some empirical tests", Studies in the theory of capital markets, ed. Michael Jensen, 79–121. New York: Praeger.

Brailsford, T./Gaunt, C./O'Brien, M. (2012): "Size and Book-to-Market Factors in Australia", *Australian Journal of Management*, 37(2).

Brückner, R. (2013): "Important Characteristics, Weaknesses and Errors in German Equity Data from Thomson Reuters Datastream and their Implications for the Size Effect", Working Paper.

Brückner, R./Lehmann, P./Stehle, R. (2012): "In Germany the CAPM is Alive and Well", Working Paper. Available at SSRN: http://ssrn.com/abstract=2161847.

Brückner, R./Stehle, R. (2012): "Der Geregelte Markt Frankfurt: Ein ökonomischer Nachruf", Working Paper.

Bühler, W./Göppl, H./Möller, H. P. (1993): "Die Deutsche Finanzdatenbank (DFDB)", *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, Sonderheft 31: Empirische Kapitalmarktforschung, 287-331.

Carhart, M. M. (1997): "On Persistence in Mutual Fund Performance", *The Journal of Finance*, 52(1), 57-82.

Chen, L./Novy-Marx, R./Zhang, L. (2011): "An Alternative Three-Factor Model", Working Paper.

Cremers, M./Petajisto, A./Zitzewitz, E. (2010): "Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation", Working Paper.

Chui, A. C. W./Titman, S./Wei, K. C. J. (2010): "Individualism and Momentum around the World", *Journal of Finance*, 65(1), 361-392.

Cuthbertson, K./Nitzsche, D. (2013): "Performance, Stock Selection and Market Timing of the German Equity Mutual Fund Industry", *Journal of Empirical Finance*, 21, 86-101.

Daniel, K./Moskowitz, T. (2013): "Momentum Crashes", Working Paper.

Daske, S./Ehrhardt, O. (2002): "Kursunterschiede und Renditen deutscher Stamm- und Vorzugsaktien", *Financial Markets and Portfolio Management*, 16(2), 179-207.

Dyck, A./Lins, K. V./Pomorski, L. (2013): "Does Active Management Pay? New International Evidence", *Review of Asset Pricing Studies*, 3(2), 200-228.

Dimopoulos, T./Wagner, H. F. (2012): "Corporate Governance and CEO Turnover Decisions", *Swiss Finance Institute Research Paper*, No. 12-16.

Elsas, R./El-Shaer, M./Theissen, E. (2003): "Beta and Returns Revisited: Evidence from the German Stock Market", *Journal of International Financial Markets, Institutions and Money*, 13(1), 1-18.

Espenlaub, S./Iqbalb, A./Stronga, N. (2009): "Datastream returns and UK open offers", *European Journal of Finance*, 15(1), 61-69.

Fama, E. F./French, K. R. (1992): "The Cross-Section of Expected Stock Returns", *Journal of Finance*, 47(2), 427-465.

Fama, E. F./French, K. R. (1993): "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33(1), 3-56.

Fama, E. F./French, K. R. (2008): "Dissecting Anomalies", *Journal of Finance*, 63(4), 1653-1678.

Fama, E. F./French, K. R. (2012): "Size, Value, and Momentum in International Stock Returns", *Journal of Financial Economics*, 105(2012), 457-472.

Gehrke, N. (1994): "Die Beziehung zwischen Markt- und Buchwerten deutscher Aktiengesellschaften: Tobins q", Wiesbaden: Deutscher Universitäts-Verlag.

Gibbons, M./Ross, S./Shanken, J. (1989): "A Test of the Efficiency of a Given Portfolio", *Econometrica*, 57(5), 1121–1152.

Gompers, P. A./Ishii, J./Metrick, A. (2010): "Extreme Governance: An Analysis of Dual-Class Firms in the United States", *Review of Financial Studies*, 23(3), 1051-1088.

Göppl, H./Schütz, H. (1995): "Die Konzeption eines Deutschen Aktienindex für Forschungszwecke (DAFOX)", Working Paper.

Gregory, A./Tharyan, R./Christidis, A. (2013): "Constructing and Testing Alternative Versions of the Fama–French and Carhart Models in the UK", *Journal of Business Finance & Accounting*, 40(1-2), 172-214.

Griffin, J. M. (2002): "Are the Fama and French Factors Global or Country Specific?", *The Review of Financial Studies*, 15(3), 783-803.

Hanauer, M./Kaserer, C./Rapp, M. S. (2012): "Risikofaktoren und Multifaktormodelle fur den deutschen Aktienmarkt", *Betriebswirtschaftliche Forschung & Praxis* (forthcoming).

Herrmann, R. (1996): "Die Karlsruher Kapitalmarktdatenbank - Bilanz und Ausblick", Working Paper.

Ince, O. S./Porter, R. B. (2006): "Individual Equity Return Data From Thomson Datastream: Handle with Care!", *Journal of Financial Research*, 29(4), 463-479.

Jensen, M. (1967): "The Performance of Mutual Funds in the Period 1945-1964", *Journal of Finance*, 23(2), 389-416.

Jagannathan, R./Wang, Z. (1996): "The Conditional CAPM and the Cross-Section of Expected Returns", *The Journal of Finance*, 51(1), 3-53.

Kan, R./Robotti, C./Shanken, J. A. (2014), "Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology", *Journal of Finance* (forthcoming).

Kiss, I./Stehle, R. (2002): "Underpricing and Long-Term Performance of Initial Public Offerings at Germany's Neuer Markt, 1997 - 2001", Working Paper.

Knez, P. J./Ready, M. J. (1997): "On the Robustness of Size and Book-to-Market in Cross-Sectional Regressions" *The Journal of Finance*, 52(4), 1355-1382.

Küting, K./Weber, C. P. (1987): "Das neue Bilanzrecht: Eine Herausforderung an die deutsche Bilanzierungspraxis", *Der Betrieb*, 40, 1-10.

Lehmann, P. (2013): "More Than Fifty Years on Stage: The Performance of German Stock Mutual Funds", published in "Three Essays on the Returns of German Stocks and Mutual Funds", Dissertation, Hamburg: Dr. Kovač.

Liew, J./Vassalou, M. (2000): "Can book-to-market, size and momentum be risk factors that predict economic growth?", *Journal of Financial Economics*, 57, 221-245.

Loughran, T. (1993): "NYSE vs NASDAQ Returns: Market Microstructure or the Poor Performance of Initial Public Offerings?", *Journal of Financial Economics*, 33(2), 241-260

Loughran, T. (1997): "Book-to-Market across Firm Size, Exchange, and Seasonality: Is There an Effect?", *Journal of Financial and Quantitative Analysis*, 32(3), 249-268.

Loughran, T./Ritter, J. R. (1995): "The New Issues Puzzle", *Journal of Finance*, 50(1), 23-51.

Michou, M./Mouselli, S./Stark, A. (2010): "Fundamental Analysis and the Modelling of Normal Returns in the UK", Working Paper.

Murphy, A./Schlag, C. (1999): "An Empirical Examination of the Effect of Dividend Taxation on Asset Pricing and Returns in Germany", *Global Finance Journal*, 10(1), 35-52.

Neuhaus, S./Schremper, R. (2003): "Langfristige Performance von Initial Public Offerings am deutschen Kapitalmarkt", *Zeitschrift für Betriebswirtschaft*, 73(5), 445-472.

Pryshchepa, O./Stehle, R. (2011): "Long-Run Stock Performance of German Initial Public Offerings: Survey and Update", Working Paper, Available at SSRN: http://ssrn.com/abstract=2039962.

Rasch, S. (1994): "Börsensegmente für Nebenwerte an Europas Börsen", ZEW Discussion Papers 94-05. Available at http://hdl.handle.net/10419/29388.

Reinganum, M. R. (1990): "Market Microstructure and Asset Pricing: An Empirical Investigation of NYSE and NASDAQ Securities", *Journal of Financial Economics*, 28(1–2), 127-147.

Ritter, J. R. (1991): "The Long-Run Performance of Initial Public Offerings", *Journal of Finance*, 46(1), 3-27.

Schlag, C./Wohlschieß, V. (1997): "Is Beta Dead? Results for the German Market", Working Paper.

Schmidt, P. S./Schrimpf, A./von Arx, U./Wagner, A. F./Ziegler, A. (2011): "On the Construction of Common Size, Value and Momentum Factors in International Stock Markets: A Guide with Applications", *Swiss Finance Institute Research Paper*, No. 10-58.

Schrimpf, A./Schröder, M./Stehle, R. (2007): "Cross-sectional Tests of Conditional Asset Pricing Models: Evidence from the German Stock Market", *European Financial Management*, 13(5), 880-907.

Schulz, A./Stehle, R. (2002): "Buchwert-Marktwert-Verhältnis, Size und Beta als Erklärungsvariable für die Renditen deutscher Aktien", Working Paper.

Shumway, T. (1997): "The Delisting Bias in CRSP Data", *Journal of Finance*, 52(1), 327-340.

Shumway, T./Warther, V. A. (1999): "The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect", *Journal of Finance*, 54(6), 2361-2379.

Stehle, R. (1994): "Eigenkapitalquoten und Fremdkapitalstruktur börsennotierter deutscher Aktiengesellschaften", *Zeitschrift für Betriebswirtschaft*, 64(7), 811-837.

Stehle, R. (1997): "Der Size-Effekt am deutschen Aktienmarkt", *Zeitschrift für Bankrecht und Bankwirtschaft*, 1997(3), 62-106.

Stehle, R./Hartmond, A. (1991): "Durchschnittsrenditen deutscher Aktien 1954-1988", *Kredit und Kapital*, 24(3), 371-411.

Viale, A. M./Kolari, J. W./Fraser, D. R.: "Common risk factors in bank stocks", *Journal of Banking & Finance*, 33(3), 464-472.

Vitols, S. (2001): "Frankfurt's Neuer Markt and the IPO explosion: is Germany on the road to Silicon Valley?", *Economy and Society*, 30(4), 553-564.

Wallmeier, M. (2000): "Determinanten erwarteter Renditen am deutschen Aktienmarkt - Eine empirische Untersuchung anhand ausgewählter Kennzahlen", *Zeitschrift für betriebswirtschaftliche Forschung*, 52(Feb.), 27-57.

Waszczuk, A. (2013): "A risk-based explanation of return patterns—Evidence from the Polish stock market", *Emerging Markets Review*, 15, 186-210.

Wheatley S./Quach B. (2013): "The Fama-French Three-Factor Model - A Report for the Energy Networks Association by NERA", NERA Economic Consulting.

Wright, S. (2004): "Measures Of Stock Market Value And Returns For The U.S. Nonfinancial Corporate Sector, 1900-2002", *Review of Income and Wealth*, 50(4), 561-584.

Ziegler, A./Eberts, E./Schröder, M./Schulz, A./Stehle, R. (2007): "Multifaktormodelle zur Erklärung deutscher Aktienrenditen: Eine empirische Analyse", *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 59(5), 355-389.

8 Appendix

A. Names of Market Segments

Until 10/2007, all exchanges had at least three segments. The *Amtlicher Markt* (until 1 July 2002 named as *Amtlicher Handel*) was traditionally the highest ("top") market segment at all exchanges. This segment was regulated by a national law (*Börsengesetz*, stock exchange act) since 1896 and supervised by government agencies. It was considered as the only "official" German market segment until May 1987.

The middle and especially the lowest segments traditionally were "only" exchange regulated. Hence, they were subject to a body of rules laid down by the local exchanges, which differed to some extent between the exchanges. They are also referred to as "non-official" market segments. Before 1987, the middle segment was named *Geregelter Freiverkehr*, after 1987 *Geregelter Markt*. The lowest segment was named *Ungeregelter Freiverkehr* before May 1988 and *Freiverkehr* afterwards. The FSE changed its name to *Open Market* in 2005. The *Open Market* was for several years subdivided into three parts: (1) the *Entry Standard* (since 2006), (2) the *First Quotation Board* (from 2008 to 2012), and (3) the *Second Quotation Board* (exists since 2008, was renamed to *Quotation Board* at the end of 2012).

Around 1995 the FSE introduced two "levels of transparency" that still exist: the *General* and *Prime Standard*. Only stocks in the *Prime Standard* are eligible to be included in the selection indices (DAX, MDAX, SDAX, and TecDAX). In 10/2007 the top and the middle segment were combined and named *Regulierter Markt*, which is a term used in the European Markets in Financial Instruments Directive (MiFID).[41] Thus, only two segments exist since 2007. However, because of the two transparency levels within the *Regulierter Markt* the FSE is practically still partitioned into three "segments".

B. German Preferred Stocks

Major differences between the German preferred stock and U.S. common stocks with inferior voting power are:

a) the German preferred stocks typically, by the company charter, have a small dividend advantage with respect to the common stocks;
b) there is no upper limit for their dividend;

---

[41] Translating the German terms is not easy in this context. The *Regulierter Markt* is often translated as 'regulated market'. However, the *Geregelter Markt* and *Amtlicher Markt* were also regulated markets. *Geregelter Markt* was often translated with 'regulated market', but this term today is mostly used to refer to the new market segment that was created in 2007.

c) typically they also have a minimum dividend, which is cumulative, that is, if it cannot be paid in one year, it must be paid in the following year(s);
d) typically they have no votes (common stocks have one vote per share); in the U.S., the common stocks with inferior voting power typically have one vote, the common stock with the superior voting power typically has 10 votes.[42]

Major differences between dual class firms in the U.S. and Germany are:

a) Dual class firms are more important in Germany than in the U.S. Gompers et al. (2010) estimate that about six percent of the publically traded companies in the United States issue more than one class of common stock. In our sample of German non-financial firms, roughly 11 % of the stocks are non-voting stocks; some of them have a very large market capitalization.
b) In Germany, in most cases, both classes are exchange listed. Typically, both classes are listed on the same exchange and in the same segment. However, cases exist in which the common stock class with voting rights is only listed in the home market and/or in the lowest segment. In the U.S., typically only the common stock class with the inferior voting power is listed.
c) In German dual class firms, typically 50 % of the shares outstanding are non-voting stocks. This is also the legal maximum; the other 50 % are the stocks with the superior voting power. In the U.S., the number of shares with superior voting power is typically a much smaller fraction of the total number of shares.
d) In the U.S., these shares are typically held by directors and managers. In Germany, usually only 50 % of these stocks are held by the majority shareholder.

C. Data Collection and Sources

C.1 Book Value of Equity

We collect book values of equity from the *Handbücher der Deutschen Aktiengesellschaft* (HBDA) for the period from 1957 to 1967. For the period from 1967 to 1990, we use the same book values of equity as Schrimpf et al. (2007) which builds on data provided by the *Deutsche Finanzdatenbank* (DFDB). From 1990 we use the Worldscope Financial Database as the primary source for the book values of equity (data type: WC03501 plus deferred taxes WC03263).

[42] For these and other details of U.S. dual class firms, see Gompers/Ishii/Metrick (2010). Whereas dual-class firms are explicitly considered in studies on IPOs, ownership and control, mergers and others, they are typically not mentioned in empirical asset pricing studies focusing on the U.S. capital market.

Both data sets cover the years until and including 1990 and consist mainly of non-consolidated annual financial statements according to the German financial reporting standards by the German Commercial Code (*Handelsgesetzbuch*, HGB). Book values of equity are adjusted for non-equity components such as subscribed capital unpaid, treasury stocks and the equity portion of special untaxed reserves.[43] The book values from Worldscope consist mainly of consolidated financial statements based on HGB (before 2005) and IFRS (after 2005). In cases in which Worldscope did not report companies' book values of equity we either use the data of Schrimpf et al. (2007), HBDA, or the *Hoppenstedt Aktienführer*.

We switch from non-consolidated statements according to the German HGB to consolidated statements, because according to Gehrke (1994) few firms published consolidated annual financial statements including foreign subsidiaries before 1986.[44] In addition, the number of firms for which we have access to non-consolidated HGB statements decreases rapidly after 2002. Consolidated statements are only applied before December 1990 if non-consolidated statements are not available. We observe that the number of firms for which Worldscope reports consolidated statements according to IFRS rapidly increases after 2001, whereas the fraction of HGB statements steadily decreases.

We make the switch in 1990 because one of our most recent data collection efforts was in the context of Brückner et al. (2012). In this study, we conduct mainly cross sectional regression analyses. Therefore, the estimated book-values of equity must be comparable at each point in time. Also, 1990 is the year that is often used to divide a long period into sub-periods.

C.2 Market Value of Equity

We generally estimate the market value of equity (firm size) as the product of the stock price and the number of shares outstanding as of the end of each month. The stock prices and the number of shares are from our own data set for the period from 1953 to 10/2007 and from Datastream (data types: UP and NOSH) for the period from November 2007 onwards. We carefully examined the quality of the data on the number of shares for the whole period using the *Hoppenstedt Aktienführer*, *Saling Aktienführer* and HBDA (fact books). In addition, we cross-checked the number of shares using our data on stock splits, stock dividends, right issues, and reverse stock splits.

We find many stocks for which the number of shares outstanding differs significantly from the number of listed shares. The most prominent example of such a firm is the Deutsche

---

[43] See Schulz/Stehle (2002) and Stehle (1994) for more details on the estimation of the book value of equity.

[44] According to Küting/Weber (1987) the *Bilanzrichtlinien-Gesetz* (BiRiLiG) from December 19th, 1985 implemented the 7th EG directive, which specified that firms have to include foreign subsidiaries in their consolidated statements for financial years starting on December 31st, 1989.

Telekom AG, which issued approximately 2.993 billion shares of which only 1 billion were listed from 11/1996 to 04/1999. We believe that adjusting the market value of equity for unlisted shares improves our firm size estimate.

Firm size is measured by the market value of the total equity of a firm. A firm's equity portfolio value is typically calculated on the basis of the common and preferred stock prices and the number of shares issued in both classes. However, many firms had, at least for some time periods, only their preferred stocks exchange listed, but not their common stocks. If only one type of stocks is exchange listed, we use its price to estimate the market value of the unlisted type.[45] We also identified few cases where only the common stocks but not the preferred stocks were listed.

We calculate the book-to-market ratio using the aforementioned aggregated market capitalization over all listed and unlisted share classes. We apply the same market value of equity as a proxy for firm size. However, we take the market value of listed shares when we calculate market-value weighted rates of return of portfolios.

C.3 Rates of Return

The data required to calculate stocks' rates of return for 1953 to 10/2007 is from our own database, primarily obtained from the *Hoppenstedt Kurstabellen*, the above mentioned fact books, the KKMDB, the *Börsenzeitung*, and Datastream. From 11/2007 onwards we only use Datastream's total return time series (data type: RI.FF) to estimate rates of return.

Our database generally contains the following data types: i) the last price of each month, ii) the number of shares outstanding, iii) dividends and information on pure stock splits, iv) stock dividends, v) right issues, vi) reverse stock splits, and vii) other financial benefits. The rates of return of firms that have multiple share classes outstanding are estimated as the value-weighted rate of return over all listed share classes.

We calculate monthly rates of return from the perspective of small domestic investors. This means that we adjust the rates of return for share reallocations from majority to minority shareholders,[46] dividends that are only distributed to minority or free shareholders[47] and corporate income tax credit (see section 2.6).

---

[45] Daske/Ehrhardt (2002) show that the prices of common stocks are on average 19.18 % higher than those of non-voting stocks (1956 to 1998). However, this difference is not stable over time; hence, we do not adjust prices of common stocks.

[46] E.g. in November 1993 FAG Kugelfischer AG, the majority shareholder of Dürkopp Adler AG, distributed one for ten shares of Dürkopp Adler AG to all minority shareholders of that company.

[47] E.g. Audi AG and MAN Roland Druckmaschinen AG.

C.4 The Risk Free Rate

We use the average one month money market rates (*Monatsgeld*) reported by Frankfurt banks as the proxy for the risk-free rate. This time-series is available on the webpage of the Deutsche Bundesbank (code BBK01.SU0104 monthly interval; BBK01.ST0104 daily interval). The monthly (daily) time-series starts in 12/1959 (1 March 1967). However, in 06/2012 the Deutsche Bundesbank stopped collecting this data. Since then we use the one-month EURIBOR (*Einmonatsgeld*) time series that are also provided by the Deutsche Bundesbank (code BBK01.SU0310 monthly interval; BBK01.ST0310 daily interval).

C.5 Data Quality

In a series of papers by e.g. Shumway (1997), Shumway/Warther (1999), Ince/Porter (2006), Espenlaub/Iqbalb/Stronga (2009) and Brückner (2013) poor data quality has been identified as an important determinant for incorrect economic inferences. Existing data sets most likely do not include all stocks listed at a time and therefore may contain a selection or survivorship bias. For small firms the probability of not being included in the data set is higher than for large firms. For surviving firms, the probability of being included is higher than for dead firms. There is also reason to believe that data quality is lower for small firms than for large firms. For our sample of German stocks listed on the FSE, we have carefully checked all ingredients of the data that goes into our rate of return calculations.

Brückner (2013) systematically compares our data set for the top segment with data from Datastream. He documents serious errors and missing stocks before 1990 and recommends not to use Datastream before 1990. There are, however, random errors after 1990 as well but those are rare and data quality seems to be sufficient.

In addition, our data set reveals that regular dividends, pure stock splits (*Nennwertum-stellungen*), rights issues (*Bezugsrechtsemissionen*) and stock dividends (*Kapitalerhöhungen aus Gesellschaftsmitteln*) contribute significantly to the rate of return of a stock in Germany. German small and large stocks differ with respect to these input factors. Because of poor data quality, empirical results might be biased.

D. Pros and Cons of Our Data Sets

We leave it to researchers and practitioners to decide which data set to use. This decision should depend on the individual setup and the underlying time period. We provide a list of arguments in favor for each data set to foster the decision process. The following arguments are in support of the first data set 'TOP':

a) The main reason for using stocks listed on the top market segment only is that empirical results are less likely affected by the large number of small and tiny firms from the lower market segments. Some empirical studies remove the smallest stocks from the final sample, e.g. Chui/Titman/Wei (2010) treat stocks as missing if their market capitalization is below the fifth percentile of all the stocks within a given country in any month.

b) A large number of U.S. studies, e.g. Ritter (1991) and Loughran/Ritter (1995), document that stocks, in the first three to five years after their IPO, underperform the market. In Germany, IPOs take place in all segments (Pryshchepa/Stehle (2011)). However, a relative large number took place in the lowest segments and in the Neuer Markt. Neuhaus/Schremper (2003) indicate a stronger long run underperformance of IPOs in the lowest segments compared to the top segment of the FSE. Therefore, empirical results for the top market segment are less likely subject to IPO effects.

c) A market microstructure effect for the NYSE and NASDAQ has been documented by Reinganum (1990). Since in the U.S. most IPOs, especially of small firms usually take place at the NASDAQ and relatively few IPOs occur at the NYSE, some argue that this is mostly an IPO effect. Loughran (1993) showed that this effect is not only related to IPOs. In the U.S., even firms from different market segments that have been listed for several years are not priced in the same way. Brückner/Stehle (2012) summarize the differences in legal supervision, admission, and listing requirements between the German market segments, which could provoke a market microstructure effect. Since most firms listed in the lowest market segments are extremely small compared to large firms from the top segment in terms of the market value of the equity, they would be primarily allocated to the lower size deciles. A market microstructure effects should therefore be considered when using the 'ALL' data set.

d) Additional reasons for not including stocks listed on the Neuer Markt apply. This market segment existed only for a few years, from 1997 to 2003. As a consequence of a large number of IPOs, nearly as many stocks were listed in this segment in 2000 as in Frankfurt's top segment. These stocks performed really well for a while, then crashed. The index level at the end was only 5 % of the maximum level in 2000. Consequently, the arithmetic and the geometric mean return for these stocks differ considerably and it is unclear, whether the standard procedures used in empirical analyses are appropriate.

Nevertheless, there are also arguments in support of the 'ALL' data set:

a) The 'ALL' data set represents in addition to the stocks listed on the top segment also the stocks listed on the middle segment and the former Neuer Markt. Therefore, one may argue that this data set provides a better coverage of the German stock market.

b) Table 2 shows that the middle segment and Neuer Markt represented a large proportion of the stocks listed at the FSE. Around 2000, the Neuer Markt also represented a large proportion of the market capitalization of the FSE.
c) Since 10/2007, the reorganization of the German stock market, only one EU regulated segment exists (*Regulierter Markt*). Stocks that were formerly listed on the middle segment or top segment now have to fulfill the same listing requirements. Brückner/Stehle (2012) argue that from 2003 onwards (following the 4th *Finanzmarktförderungsgesetz*, FMFG) the listing requirements of the middle segment are similar to those of the top segment. This questions to continue looking at the former top segment, *Amtlicher Markt,* only, especially after 10/2007.

E. Daily Data

We also provide daily data of Fama/French factor sets for Germany. The daily data sets are only available from 1990 onwards and are based on the same raw data as used for the monthly data. Precisely, the process of sorting and portfolio formation is exactly the same as well as the underlying data. The only exception is the daily return data of individual stocks. The daily returns are based on Datastream's total return index (RI). We therefore cannot provide data with the corporate income tax credit. Also, the provided daily data may not accomplish the same data quality as we promise for our monthly data, especially for the time period before 2001.

F. Construction of the Market Portfolio, the Size, Value and Momentum Factors

F.1 Market Portfolio

For the construction of the market portfolio, each firm's return for month m is weighed by its market value of all listed shares from the end of month $m-1$. The time-series data is based on all stocks, including financials, and contains all financial benefits to stockholders discussed in this paper.

F.2 Size, Value and Momentum

We follow Fama/French (1993) to construct the *SMB* (small minus big) and *HML* (high minus low) factor-mimicking portfolios. For the construction of the Carhart (1997) *MOM* (winner minus loser) factor we follow Fama/French (2012).

We form portfolios based on firm's size, value and momentum starting at the end of June 1958. We only assign firms to our portfolios if they have:

i) A positive book value of equity (BE) and market value of equity (ME) as of December in year $t-1$ to estimate the book-to-market (BE/ME) ratio;
ii) ME at the end of June in year t of at least five mln., or if less, a stock price of at least one;
iii) A return history of at least 12 month to estimate momentum.

Non-financial firms that fulfill these requirements are independently sorted according to their ME, BE/ME, and momentum. The median of the ME values (June, year t) is the intersection for the size sorting. All firms with ME larger than the ME median are big (B), others are small (S). The value sorting is based on the book-to-market ratios as of December in year $t-1$. The 0.3 and 0.7 quantiles of BE/ME values are the intersections for the value sorting. Thus, we construct three groups that split firms in low (L), medium (M) and high (H) book-to-market firms. In this way, we construct six double-sorted portfolios: S/L, S/M, S/H and B/L, B/M, B/H. Each firm is, according to his individual ME and BE/ME value, allocated to one of those portfolios and remains there until June next year or until its delisting.

In contrast, the size/momentum portfolios are formed on a monthly interval. Each month m, we calculate stock's past return (momentum return) from $m-2$ to $m-12$ (11 month, skipping the most recent month).[48] We then range all stocks according to their past performance and extract the 0.3 and 0.7 quantile. Firms that performed bad in the past are called losers (Lo) and are part of the first group, firms with normal (N) performance (they make up 40 % of all firms) are in the second group and past winners (W) are in the third group. Based on firm's individual momentum and ME from last June, we allocate each observation to one of the six double-sorted portfolios: S/Lo, S/N, S/W and B/Lo, B/N, B/W, where, e.g., portfolio B/W represents big winners.

In a second step, we estimate value-weighted rates of return for the above mentioned portfolios. We weight a firm's return for month m by its market value of all listed shares from the end of month $m-1$. To avoid any selection bias we remove only firms that are delisted during the period from July in year t to June in year $t+1$. In a final step, we calculate the SMB, HML and MOM factors as

$$SMB_m = \frac{1}{3}\left(S/L_m + S/M_m + S/H_m\right) - \frac{1}{3}\left(B/L_m + B/M_m + B/H_m\right), \quad (2)$$

[48] Formation date is the end of month $m-1$, see e.g. Figure 1 in Daniel/Moskowitz (2013) for a graphical illustration.

$$HML_m = \frac{1}{2}\left(S/H_m + B/H_m\right) - \frac{1}{2}\left(S/L_m + B/L_m\right) \quad (3)$$

and

$$WML_m = \frac{1}{2}\left(S/W_m \; + B/W_m \;\right) - \frac{1}{2}\left(S/Lo_m + B/Lo_m\right). \quad (4)$$

G. Construction of Fama/French Factors based on MSCI Indices

Using MSCI indices is a convenient alternative to calculate Fama/French factors since the data for the relevant indices are available online. We obtain the “gross” (total) return time series in "local” currency and estimate the return on the market portfolio, SMB and HML factor portfolios based on the following indices. The rate of return on the market is based on the MSCI Germany. The SMB factor is the difference between the rates of returns on the MSCI Germany SMID Cap (small and mid caps) and the MSCI Germany Large Cap. For the HML factor, we use the MSCI Germany Value and MSCI Germany Growth indices. Unfortunately, MSCI does not provide a momentum index for Germany, only a European index seems to be available. Important to note, the calculation of the growth and value indices are not comparable to the methodology suggested by Fama/French. E.g., the value index by MSCI also takes forward looking variables as the “12-month forward earnings to price ratio” into account.[49]

[49] For details see http://www.msci.com/products/indices/style/methodology.html.