Estimating Beta-Coefficients of German Stock Data: A Non-Parametric Approach

Maik Eisenbeiß (maik.eisenbeiss@uni-muenster.de)
Westfälische Wilhelms-Universität Münster

Göran Kauermann (gkauermann@wiwi.uni-bielefeld.de)
Universität Bielefeld

Willi Semmler (semmlerw@newschool.edu)
Universität Bielefeld and New School University New York

Abstract. Although the consumption based asset pricing theory appears to be theoretically superior and more elegant than the beta pricing model, yet in practice the beta pricing model is more widely applied. Indeed, beta pricing models are one of the most widely adopted tools in financial analysis. They easily allow to handle systematic risk as priced in financial assets. However, accurately estimating beta-coefficients is not as straightforward as implicitly suggested by Sharpe's standard market model, i.e., simply using the ordinary least-squares (OLS) regression. This is primarily because beta-coefficients cannot generally be assumed as being stable over time. In order to overcome this deficiency, we present and apply a non-parametric estimation technique that allows capturing this time effect and promises both, more reliable estimates than obtained with an OLS-regression as well as a better manageability compared to the existing econometric approaches dealing with time-varying beta-coefficients. Estimation results for constant and time varying betas are presented for portfolios of German industries.

Keywords: systematic risk, time-varying beta-coefficients, non-parametric estimation, varying-coefficient model

1 Introduction

In the financial literature there are two paradigms used for studying asset pricing and portfolio decisions. The first paradigm is static and is based on the traditional capital asset pricing model (CAPM) of Markowitz (1952) and Sharpe (1964), whereas the second approach is based on intertemporal decisions of economic agents who exhibit well specified preferences for consumption over time. The latter approach is called consumption based asset pricing which is well grounded in economic theory. It answers most asset pricing and portfolio decisions in principle, yet it does not work well in practice. The latter attempts to tie the stochastic discount factor, crucial for pricing assets, to preferences and consumption data, but its performance in practice is rather limited.[1]

On the other hand, models that use factors for asset pricing are easy to handle in practice since they are usually obtained in linear form through linear regressions. The typical example for this is the CAPM where a beta, as a price for risk, prices the asset and justifies the return from an asset or portfolio of assets. The CAPM also allows to obtain a discount factor in linear form from the beta estimations[2]. If the presumed consumption based asset pricing model is of special form, for example representing special preferences then the CAPM represents an equivalent form of asset pricing[3].

Apart from the ongoing research on the theoretical relationship of the CAPM to the consumption based asset pricing model, in practice beta estimates have always been used as a measure of risk in modern finance. They have contributed to a variety of applications such as testing of asset pricing theories, estimating cost of capital, hedging market exposure as well as portfolio performance evaluation and thus they are one of the most widely adopted instruments among practitioners and financial economists in order to measure and manage risk. Wells (1995, p. 5) nicely summarizes the significance of beta in just one sentence: "It is one of the few regression coefficients, simple or otherwise, that people actually pay money to get." As a consequence, accuracy in the measurement of the beta-coefficient can be considered a striking topic.

[1] For an extensive evaluation of those two theories, see Cochrane (2001, ch. 9) and Campbell and Cochrane (2000).

[2] See Cochrane (2001, chs. 6 and 9).

[3] See Cochrane (2001, ch. 9).

During the last three decades numerous studies have addressed the question of beta's stability over time. Among the most prominent studies are Blume (1971), Baesel (1971), Altman, Jacquillat, and Levasseur (1974), Roenfeldt, Griepentrong, and Pflaum (1978), Alexander and Chervany (1980) and Theobald (1981). What they share in common is the observation that beta-coefficients are far from being stable. Despite this general consensus, still the most widely adopted approach for estimating beta is the ordinary least-squares regression. In its simplest form it is often based on the following model framework, known as the *market model*:

$$R_j = \alpha_j + \beta_j R_M + \varepsilon_j \tag{1}$$

with R_j and R_M denoting the holding period excess return on the jth security and on the overall market M, respectively[4]. The parameter α_j depicts the security's expected return if the market is neutral, i.e., if $R_M = 0$, and ε_j quantifies changes in the holding period excess return due to changes that are exclusively firm-specific. The beta-coefficient is, of course, depicted by β_j and assesses the impact of movements in the market on the jth security. However, such an approach assumes constant coefficients and in the context of time-varying betas it is likely to produce inconsistent results. One reason for its prevailing use might be due to the fact that its usage is relatively straightforward and that the existing econometric alternatives, which promise more accurate estimates, are much harder to pursue. Among those econometric alternatives the most widely adopted approach is to estimate beta as a time-series process using the *Kalman filter*. In this respect, the time-path of beta is commonly modelled as a process which relates today's beta value either to its overall mean or to last periods' beta values[5].

The remainder of this paper concentrates on a fundamentally different approach: Using a non-parametric estimation technique, we treat the beta-coefficient as an unspecified function of time. The estimation results which are based on daily data of industry replicating portfolios for the German stock market between 1992 and 2003 support our approach as a useful and especially easy accessible compromise in the environment of estimating time-varying beta-coefficients. Moreover, the results do not only confirm the

[4]The holding period excess return depicts the holding period return in excess of the risk-free rate r_f.

[5]See Schaefer, Brealey, Hodges, and Thomas (1975)

general hypothesis of non-constant betas but they also suggest a counter-cyclical behavior of several industry groups' beta-coefficients with respect to the state of the market. On average, betas tend to be significantly larger in bear-markets than in bull-markets for these industry groups. Hence, market phases can be considered a driving force for the variability of beta-coefficients.

The paper is structured as follows: Chapter 2 introduces the non-parametric approach for estimating time-varying beta-coefficients. Chapter 3 then provides a overview of the data before discussing the main results. Chapter 4 finally draws the main conclusions.

2 A Non-Parametric Estimation Approach

Essentially, the starting point of the non-parametric estimation approach is to generalize the market model (1) by including a component which accounts for the time-variation in the beta-coefficient. The "generalized" market model is given by

$$R_{jt} = \alpha_j + \beta_j R_{Mt} + \tilde{\beta}_j(t) R_{Mt} + \varepsilon_{jt} \tag{2}$$

with $\tilde{\beta}_j(t)$ capturing the time-effect embedded in the systematic risk component of the jth security. Apparently, coefficients β_j and $\tilde{\beta}_j(t)$ in (2) are not identifiable unless we impose $\int \tilde{\beta}_j(t)dt = 0$. In this respect, the parametric beta-estimate can be regarded as the mean value of beta for the whole period under consideration. Thus, dependent on the time-period it is additively increased or decreased by the estimate of its non-parametric counterpart. In other words, the estimated beta-coefficient in period k is given by $\hat{\beta}_k = \hat{\beta} + \hat{\tilde{\beta}}(k)$.

In the context of the non-parametric estimation literature, a model of this form is known as a *varying-coefficient model*[6]. The main idea of a varying-coefficient model is to provide a framework in which the influence of the predictor variables on the response variable is linear while the corresponding coefficients are no longer treated as constants but rather as functions of other variables. In particular, $\tilde{\beta}_j(t)$ expresses the multiplicative interaction or temporal changes, respectively, of the influence of R_{Mt}.

[6]The varying-coefficient model was first proposed by Hastie and Tibshirani (1993).

A varying-coefficient model as in (2) can be fitted to data by finding the solution of the following penalized least-squares criterion[7]

$$\min_{\alpha_j,\beta_j} \quad \sum_{t=1}^{T} \left(R_{jt} - \alpha_j - \beta_j R_{Mt} - \tilde{\beta}_j(t) R_{Mt} \right)^2 + \lambda_j \int_a^b \tilde{\beta}_j''(t)^2 dt. \qquad (3)$$

While the first term measures the goodness of fit, the second term penalizes the curvature of the function $\tilde{\beta}_j(.)$. The so-called *smoothing parameter* λ_j deserves special attention, since it controls the trade-off between bias and variance of the fit. For $\lambda \to 0$ the influence of the penalty term disappears and the resulting function tends to interpolate between the observations. In contrast, $\lambda \to \infty$ forces the penalty term to dominate and thus yields a simple linear regression fit. In this respect, Wood (2000) has proposed an algorithm which "automatically" determines an "optimal" level of smoothing. This algorithm is provided as the function `gam()` in the package `mgcv` of the (public domain) statistical software environment R. The appropriate R-commands can be found in Appendix A.

In (2) we have not said much about the structure of the residuals ε_{jt}. The natural assumption of homoscedasticity is likely to be too simplistic and two potential violations spring in our mind: First, residuals can be correlated and secondly, the residual variance can change over time. In the first case, the automatic smoothing parameter selection method by Wood (2000) would fail in the presence of autocorrelated errors. In such a case, the smoothing parameter should be determined by hand or by other more elaborated methods suggested in the paper by Opsomer, Wang, and Yang (2001). In the data example at hand autocorrelation was not observable, based on both, a graphical investigation and the Durbin Watson statistics.

In contrast, we found clear indication of heteroscedasticity in the data after investigation of the fitted residuals based on a first smooth estimate. Note that homoscedasticity is an integral assumption in the estimation step as well as when drawing inference from the estimation results. Hence, in the case of heteroscedastic residuals the estimation approach as described above does not necessarily guarantee reliable or at least efficient estimates. We cope with heteroscedastic residuals by pursuing a two-step estimation approach: The first step is to determine the squared residuals from the

[7]A technical introduction into a fitting mechanism of this kind for a varying-coefficient model can be found in Appendix C.

estimation results of model (2) by assuming (working) homoscedasticity. This yields working residuals defined through

$$\hat{\varepsilon}_{jt} = (R_{jt} - \hat{R}_{jt})^2. \tag{4}$$

A simple exploratory investigation is available by plotting $\hat{\varepsilon}^2_{jt}$ against t. Heteroscedasticity is now visualized by the structure in the plot. In a non-parametric and flexible way we can model dependence of residual variation on time with a *generalized additive gamma model*. This is accommodated by modelling the squared residuals as

$$\mu^2_{jt} = Var(\varepsilon^2_{jt}) = E(\varepsilon^2_{jt}) = g\{\alpha_j + f_j(t)\} \tag{5}$$

where $f_j(.)$ is a smooth but otherwise unspecified non-parametric function and $g(.)$ is called the inverse link function. Based on a gamma model we choose $g^{-1}(\mu) = -1/\mu$.[8] The major idea behind the gamma model is, that the variance of ε^2_{jt} is proportional to μ^4_{jt} and with $f_j(t) = constant = 0$ a homoscedastic (normal residual) model results. Hence, $f_j(t)$ captures the heteroscedasticity over time. Model (5) falls also in the class of models which can be fitted with the `gam(.)` procedure in R as demonstrated in the appendix. In particular, model (5) can be regarded as a special case of the family of *generalized additive models*, which themselves are a generalization of additive regression models[9]. The scedasticity-structure estimated by the gamma-model can now be used to refit model (2) in a weighted manner which implicitly accounts for heteroscedasticity. Let therefore $\omega_{jt} = 1/\mu^2_{jt}$ be weights constructed from the fitted model (5). These weights are now used to fit model (2) in a weighted form. In practice, this can be pursued by inserting weights in the fitting routine, as it is implemented in the so far used R software. More details are found in Appendix A.

[8] See McCullagh and Nelder (1989).

[9] See Hastie and Tibshirani (1990). While additive models *linearly* associate the response variable with an additive sum of (non-parametric) functions of the predictor variables, generalized additive models allow the response variable to depend on the additive predictor through a *nonlinear* relationship.

3 Data and Results

3.1 The Data

In this section we want to demonstrate the potentials of the non-parametric estimation approach by applying the model to a variety of industry groups based on data from the German stock market.

The data used are daily observations from April 1991 to March 2003. In its original form, it comprises stock prices for a range of selected companies with listed securities at the German stock market as well as observations on the CDAX performance index. The data are from the Reuters 3000 Xtra database. Concerning the risk-free rate in the market model we use overnight money market rates obtained from the German Bundesbank[10].

The following seven industry groups will be regarded in the analysis[11]: Automobile, Banks, Consumer, Industrial, Retail, Pharma & Healthcare and Utilities. The most obvious and, of course, the most accurate way to refer to these industry groups would be to consider several industry-specific stock market indexes. However, due to lacking data, the approach pursued in this analysis is to construct industry-replicating portfolios containing major companies of the specific lines of business based on a capitalization weighting method[12]. This approach is justified by the fact that companies within an industry group can be assumed to share several common characteristics such as their sensitivity to business cycles, international tariffs, technological development or raw material availability. Hence, the beta-risk of an industry-replicating portfolio approximates the beta-risk borne by the whole industry.

The stock prices were converted to discrete rates of return[13]. In the context of this analysis, discrete returns are preferable over continuously compounded returns since they retain the property of *additivity within portfolios.*

As suggested by the market model, a broad market index should be used in order to approximate the effects of common macroeconomic events. For the German stock market, the CDAX performance index can be regarded as

[10] Data-code: ST0101

[11] The industry groups are defined corresponding to the sector indexes of the prime segment as described in Deutsche Börse Group (2003).

[12] Table B.3 in appendix B provides a summary of the companies included in the analysis.

[13] Dividend payments can be disregarded, since the focus is on daily returns.

an adequate proxy for the common macro factor. Strictly speaking, it contains all *domestic* listings of the stock market segments *Prime Standard* and *General Standard*. These segments represent the entire range of both domestic and foreign securities listed at the German stock market. In analogy to the stock prices, the index data was also converted to discrete returns.

The overnight money market rates, taken as a proxy for the risk-free rate in the market model, are quoted as annual rates of return based on the act/360 standard. This method annualizes the rates of return assuming a 360-day year. Thus, in order to obtain daily rates the annual yields were simply divided by 360.

3.2 Empirical Results

As outlined above, the non-parametric estimation results based on the generalized market model (2) can only be considered reliable as long as the residuals are neither autocorrelated nor heteroscedastic: Figure B.4 shows the autocorrelation function of the fitted residuals in the different industry groups, with model (2) fitted under the assumption of uncorrelated errors. Apparently, there is no indication of autocorrelation which is also shown in the Durbin Watson statistics provided in Table B.4. Contrary, when looking for homoscedasticity Figure B.5 with fitted residuals $\hat{\varepsilon}_{jt}^2$ plotted against time clearly gives the impression that the validity of homoscedasticity is questionable. For this reason, the two-step estimation routine as suggested in the previous section will be pursued.

The weighted estimation diagnostics for the various industry groups are reported in Table 3.1. As seen, both the parametric estimates reflecting the mean value of systematic risk for the whole period under consideration as well as the smooth terms are statistically significant at the 99%-level for each industry group. The mean value of systematic risk of the industry groups "Automobile" and "Banks" is close to unity, implying that, on average, these groups are affected by macroeconomic events to the same extent as the average market. In contrast, all other groups are less responsive to overall market movements. The smallest parametric beta-estimate of only 0.55606 can be observed for the "Pharma & Healthcare" industry portfolio. Interestingly, on average there is no industry group that "overreacts" to macroeconomic events, meaning that its parametric beta-estimate turns out

Industry Group	Parametric Term			Smooth Term	
	Coeff.	Std.Error	P-Value	EDF	P-Value
Automobile	0.98661	0.02348	< 0.001	16.71	< 0.001
Banks	0.93439	0.02079	< 0.001	14.55	< 0.001
Consumer	0.59742	0.02441	< 0.001	6.18	< 0.001
Industrial	0.77839	0.02119	< 0.001	15.45	< 0.001
Retail	0.67999	0.02279	< 0.001	4.37	0.0016
Pharma & H.	0.55606	0.02332	< 0.001	4.32	< 0.001
Utilities	0.73886	0.02155	< 0.001	15.95	< 0.001

Table 3.1: Weighted estimation results for the different industry groups

to be greater than one.

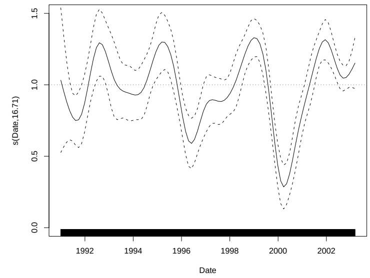
However, as illustrated in Figure 3.1, the estimated beta-coefficients are far from being stable over time and significantly deviate from their respective mean values, especially during the period of 1996 – 2002. To be more precise, the estimation results roughly subdivide the industry groups into two classes as far as the volatility of the time-path of systematic risk is concerned:

One class comprises the industry groups "Automobile", "Banks", "Industrial" and "Utilities" with each having a (comparably) highly volatile beta-coefficient. In contrast, the various beta-coefficients of the industry groups "Consumer", "Retail" as well as "Pharma & Healthcare" are exposed to a comparably small degree of time-variation. Strictly speaking, the estimated beta-paths "smoothly" decrease over time with troughs around the years 2000 and 2001.

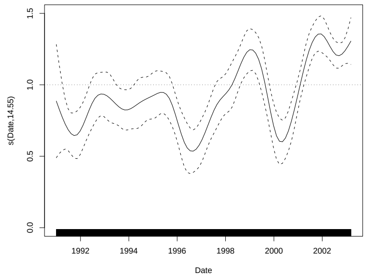
Especially for the industry groups with highly volatile beta-coefficients, a relationship between the time-path of systematic risk and the overall market conditions can be seen. Figure 3.2 captures the return on the CDAX as a smooth function of time[14]. Denoting periods of average returns greater than zero as up- or bull markets, and periods of average returns smaller than zero as down- or bear markets, the following pattern can be observed: *For these industry groups beta-coefficients tend to move counter-cyclically with respect to the state of the market: on average, the respective beta-coefficients are large in bear markets and small in bull markets.* This also means that investors focussing on these groups receive a higher risk-premium in periods of substantial downside variation than in periods of upside variation.

More precisely, during the bearish market phases A, C and E as de-

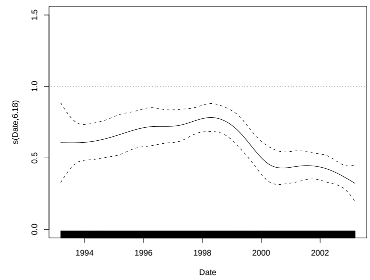
[14] The smooth function is based on an estimated gam-model of the form $r_{Mt} = f(t) + \varepsilon_t$ using 14 degrees of freedom.



Automobile
Maximum basis dimension: k=20

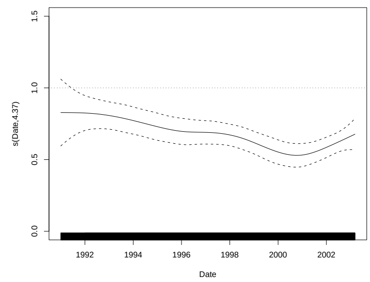


Banks
Maximum basis dimension: k=20

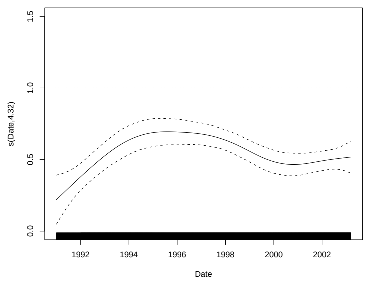


Consumer
Maximum basis dimension: k=10

Industrial
Maximum basis dimension: k=20



Retail
Maximum basis dimension: k=10



Pharma & Healthcare
Maximum basis dimension: k=10

Utilities
Maximum basis dimension: k=20

Figure 3.1: Estimated time-paths of $\beta_j + \tilde{\beta}_j(t)$ for the various industry groups

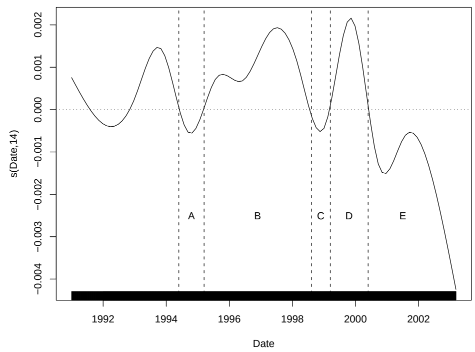


Figure 3.2: Trend of the returns on the CDAX performance index

picted in Figure 3.2, the respective beta-coefficients of the industry groups "Automobile", "Banks", "Industrial" and "Utilities" are large or, at least, exhibit an upward trend. Conversely, the bullish market phases B and D are associated with smaller beta-coefficients as far as these industry groups are concerned.

3.3 Focussed Estimates

In order to cross-validate the suggested counter-cyclical behavior of beta for these industry groups, it seems helpful to focus on a reduced data-range. This is advisable since the estimation results are based on "global" concepts, that is to say smoothness is refined to as smoothness over the full range of time[15]. As a consequence, local structures might be appropriately exhibited. Hence, we subsequently concentrate on a reduced estimation interval covering the market phases C, D and E as denoted in Figure 3.2.

Considering the various time-paths of systematic risk for the industry groups "Automobile", "Banks", "Industrial" and "Utilities" in the reduced estimation interval in Figure 3.3, the suggested counter-cyclical behavior of the beta-coefficient becomes even more evident. All of them exhibit a small or at least decreasing beta-coefficient during the bullish market phase D and a large or increasing beta-coefficient during the bear market environments C and E, respectively. Thus, the estimation results based on the reduced

[15] Spline smoothing in contrast to local smoothing is a global optimization problem. See Hastie and Tibshirani (1990)

sample support the suggestion of a counter-cyclical behavior of the beta-coefficient for these industries.

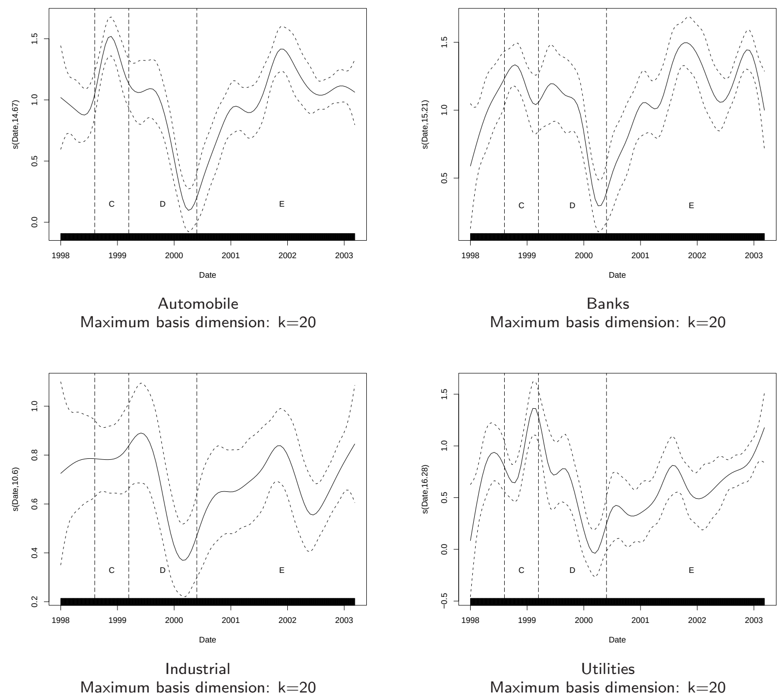


Figure 3.3: Estimated time-paths of $\beta_j + \tilde{\beta}_j(t)$ between 1998 and 2003

Of course, since the beta-coefficient of the market (portfolio) is, by definition, equal to one, there must also be some industry groups that exhibit the exact opposite behavior. However, as far as our industry sample is concerned, there is no evidence that such industry groups are included.

3.4 Performance Comparison

Finally, one aspect remains worth addressing: Even though the non-parametric estimates as presented so far reveal several "neat" results and insights into the time-path of systematic risk and thus indirectly acknowledge the non-parametrical estimation technique as a useful tool for estimating beta-coefficients, the real gains in estimation accuracy still remain to be unknown. Therefore, this section aims to *quantify* the realized gains in estimation accuracy by measuring the improvements in fit over the conventional ordinary least-squares approach for constant beta-estimates. Generally, two measures should be compared: first, the adjusted R^2 as a measure of overall fit, and secondly the Generalized Cross Validation (GCV)-criterion. The GCV-criterion approximates the mean-squared error and hence accounts for both bias and variance of an estimate. The objective is to reduce the GCV function[16].

The results are reported in Table 3.2. Overall, they show gains in accuracy for each industry group in terms of a higher R^2 and lower GCV-score when beta is allowed to vary. As expected, the largest improvements can be observed for those industry groups whose betas are exposed to a high degree of time-variation, especially for the industries “Automobile”, “Banks” and “Utilities”. Thus, the non-parametric estimation technique as applied in this analysis can be considered a useful tool for estimating time-varying beta-coefficients. Moreover, it provides a comprehensible alternative compared to the existing time-series approaches for these estimation purposes.

Industry Group	**Constant Beta**		**Time-Varying Beta**	
	R^2(adj)	GCV	R^2(adj)	GCV
Automobile	0.525	1.03230	0.564	0.97797
Banks	0.550	1.04040	0.583	0.98471
Consumer	0.202	1.01920	0.225	0.99640
Industrial	0.377	1.01440	0.391	0.99495
Retail	0.278	0.99793	0.284	0.99398
Pharma & H.	0.210	0.99891	0.215	0.99100
Utilities	0.320	1.01610	0.341	0.99893

Table 3.2: Performance comparison between constant and time-varying beta-estimates

4 Conclusion

Our empirical results reveal that the German stock market exhibits symptoms of time-varying beta-coefficients. This insight is in accordance with previous studies that have found evidence of beta-instability in various other countries. However, this study differs from preceding research in the method applied in order to estimate the time-path of beta. While most studies have estimated beta as a time-series process using the Kalman filter, we focus on a non-parametric estimation technique that allows to treat the beta-coefficient

[16] See also Appendix C.

as an unspecified function of time. Compared to the existing time-series models, such an approach is not only more intuitive and thus quite easy to understand, the corresponding model can also be fitted to data in a comfortable manner, as shown in Appendix A. Moreover, heteroscedasticity is easily accommodated in the same model framework.

Due to the flexibility of the model by treating the beta-coefficient as an unspecified function of time, we were able to identify several industry groups as being counter-cyclically related to the state of the market as far as their systematic risk exposure is concerned: on average, betas tend to be significantly larger in bear-markets than in bull-markets for the groups "Automobile", "Banks", "Industrial" as well as "Utilities".

With respect to the importance of the beta-coefficient in modern finance, the results of this paper might be helpful in two regards. On the one hand, non-parametric estimation techniques whose real strengths are still underrated in many economic disciplines are shown to provide a potential alternative to the more complicated time-series approaches for estimating time-varying beta-coefficients. On the other hand, the empirical analysis reveals significant insight into the time-path of systematic risk for a variety of industry groups at the German stock market. Among the most important finding is the counter-cyclical behavior of some industry groups' beta coefficients with respect to the state of the market. For instance, with knowledge of this kind beta-predictions could simply be used to adjust for the expectations of future market conditions.

A R-Commands

Software package R is an open source software which can be downloaded free of charge from http://www.r-project.org. The routines used in this paper are implemented in the mgcv package also available from the above web page. A general overview about the features of R is found for instance in Dalgaard (2002).

The generalized market model (2) can be estimated using the following R-command:

```
> gam(Rj ~ RM + s(t, by=RM))
```

The by-argument ensures that the smooth function $\tilde{\beta}_j(t)$ gets multiplied by the predictor R_{Mt}. The formula further indicates that the estimated time-path of the beta-coefficient is represented by a constant plus a smooth effect. This is because the individual smooth functions in a varying-coefficient model have to be constrained to have zero mean. Otherwise the effects of the covariates would not be identifiable. Since mgcv's gam() automatically accounts for this constraint, regardless of the number of covariates in the model, the resulting smooth functions are centered around zero. However, for model (2) this means that the estimate of the whole term $\beta_j(t)R_{Mt}$ is centered around zero, even though its "actual" mean value might differ from zero. Therefore, the term R_{Mt} must be included as well whose parametric estimate can be regarded as the mean value of beta for the whole period under consideration. Thus, depending on the time period the smooth component either increases or reduces the estimated mean value of the beta-coefficient.

A.1 Autocorrelated Residuals

One way to cope with autocorrelated residuals would be to determine the smoothing parameter by hand. In such a case, the appropriate gam()-formula is as follows:

```
> gam(Rj ~ RM + s(t, by=RM, knots|f))
```

The parameter knots must be replaced by the desired number of knots: the more knots are placed the more flexible the fit becomes and vice versa.

A.2 Heteroscedasticity

The two-step estimation approach for coping with heteroscedastic residuals first requires to fit the generalized additive gamma model (5). In R this can be done using the following command:

```
> gam(r.squared ~ s(t), family=Gamma)
```

with the `family`-argument specifying the desired response probability distribution. After having constructed the appropriate weights, the weighted market model can be fitted to data by:

```
> gam(Rj ~ RM + s(t, by=RM), weights=wj)
```

with the `weights`-argument ensuring to incorporate the vector of the constructed weights w_{jt}.

B Data Summary and Estimation Diagnostics

Company	1st Obs.	Company	1st Obs.
Automobile		**Banks**	
Volkswagen AG	04/1991	Bayer. H.- und Vereinsbank AG	04/1991
Daimler Chrysler AG	04/1991	Commerzbank AG	04/1991
BMW AG	03/1993	Deutsche Bank AG	04/1991
Continental AG	03/1993	IKB Dt. Industriebank AG	01/1996
Consumer		**Retail**	
Adidas-Salomon AG	11/1995	Celesio AG	01/1996
Henkel KGaA	03/1993	Douglas Holding AG	04/1995
Wella AG	01/1996	Fielmann AG	01/1996
Beiersdorf AG	11/1996	Karstadt Quelle AG	04/1991
Puma AG	07/1996	Metro AG	07/1998
Pharma & Healthcare		**Industrial**	
Altana AG	05/1996	Deutz AG	11/1992
Fresenius Medical Care AG	10/1996	Linde AG	04/1991
Schering AG	06/1991	MAN AG	04/1991
Merck KGaA	10/1995	Rheinmetall AG	03/1998
Schwarz Pharma AG	02/1996	MG Technologies AG	10/1992
		Thyssen Krupp AG	04/1991
Utilities		IWKA AG	08/1996
E.ON AG	03/1993	Jenoptik AG	10/1998
RWE AG	04/1991		

Table B.3: Companies included in the analysis with dates of their first observation

Industry Group	DW-statistic
Automobile	2.0167
Banks	2.0758
Consumer	2.1880
Industrial	2.0186
Retail	2.0102
Pharma & Healthcare	1.9751
Utilities	1.9457

Table B.4: Durbin Watson statistics for the various industry groups.

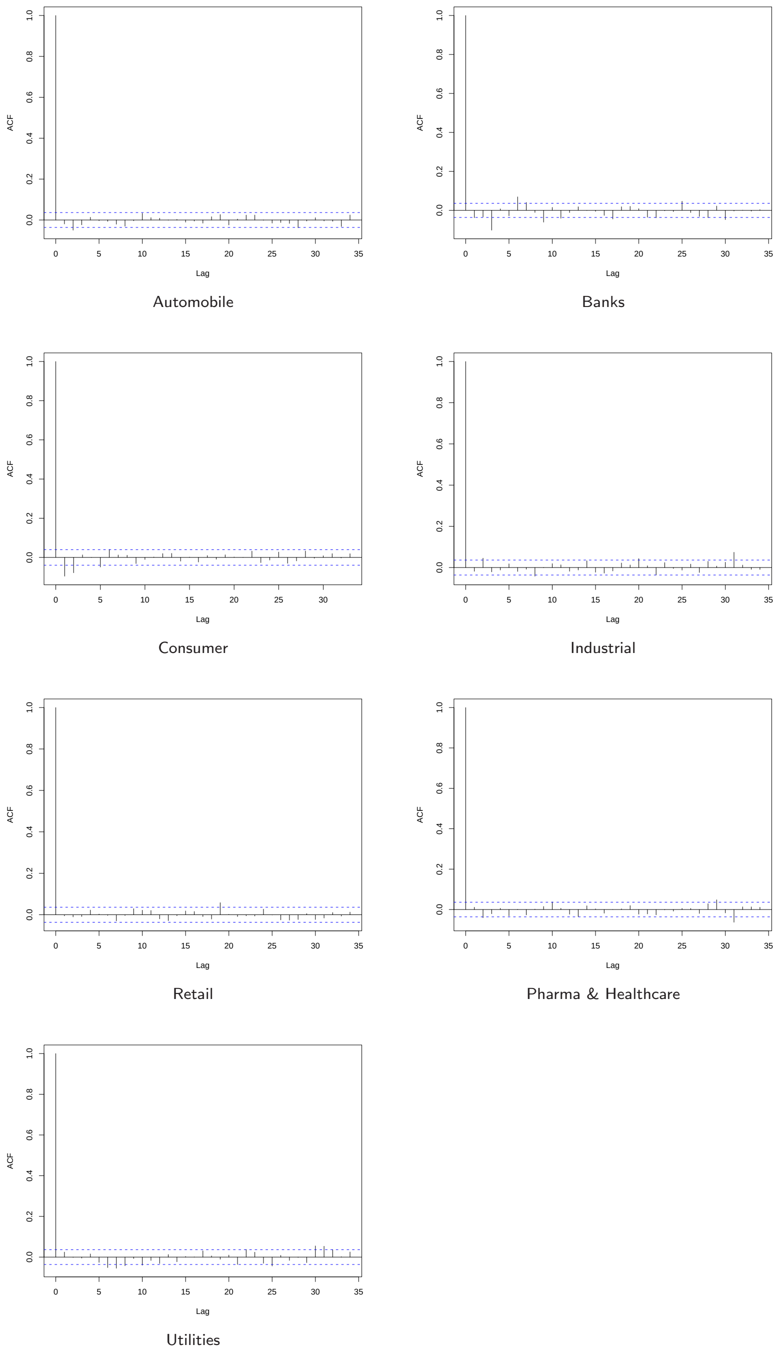


Figure B.4: Unweighted estimation: ACF for the various industry groups.

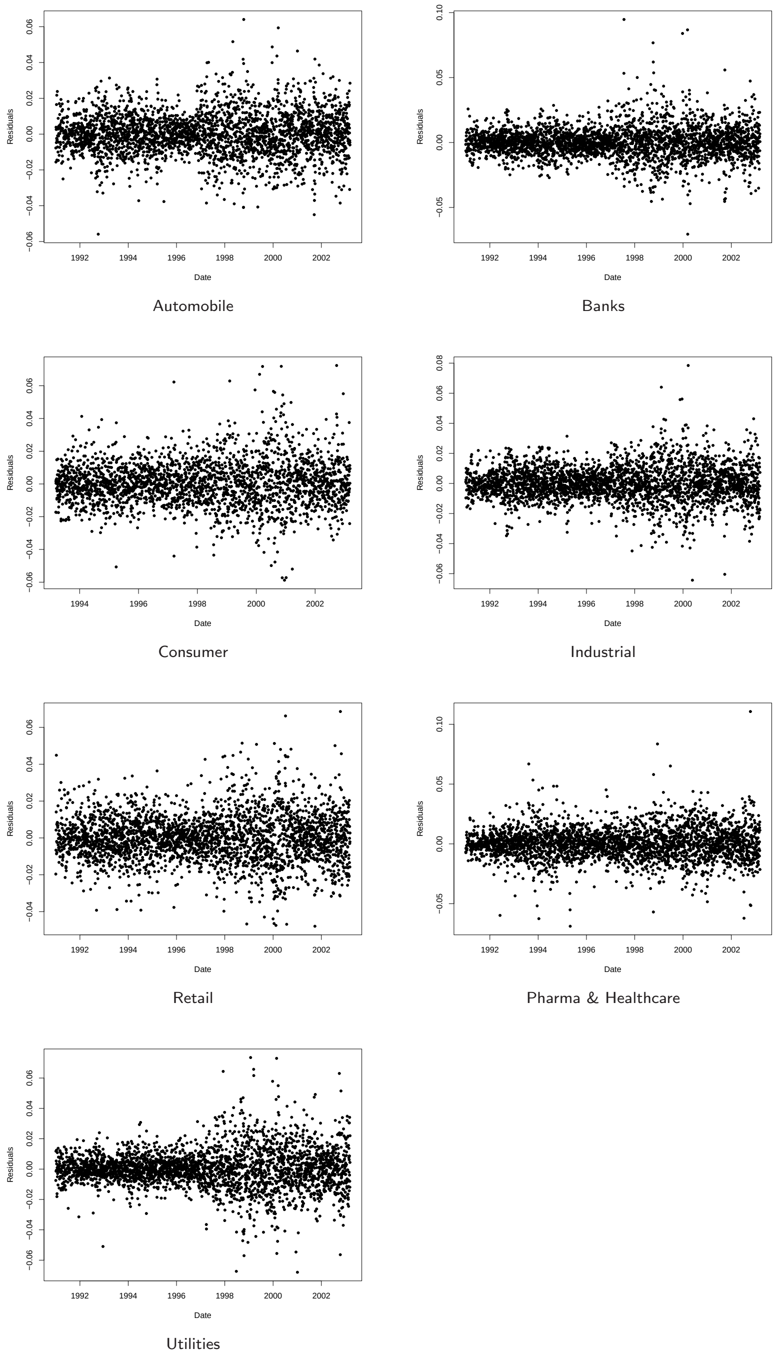


Figure B.5: Unweighted estimation: Residuals for the various industry groups.

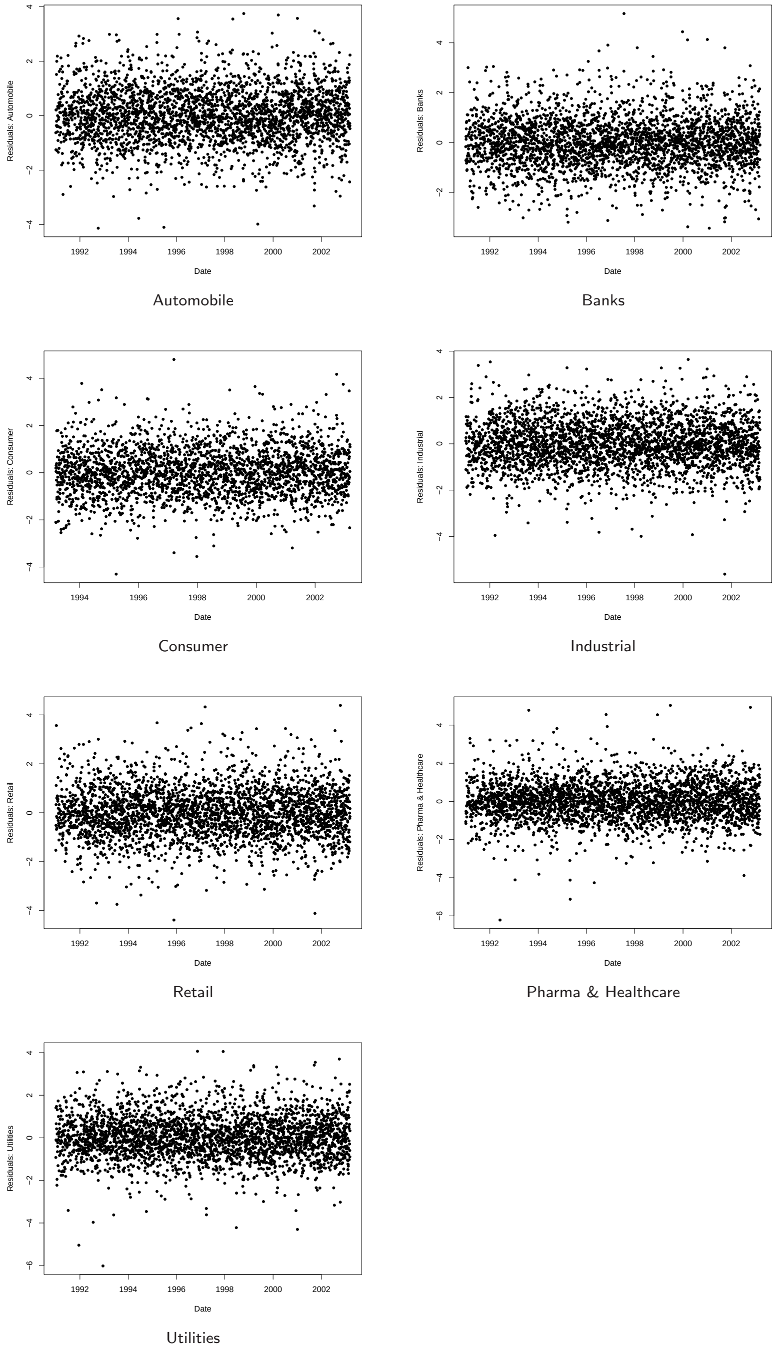


Figure B.6: Two-step weighted estimation: Residuals for the various industry groups.

C Theoretical Notes on the Varying-Coefficient Model

It can be shown (see e.g. De Boor (1978, ch. 4)) that a weighted version of the minimization problem (3) can be written as

$$(\boldsymbol{Y} - \boldsymbol{X}\beta - \boldsymbol{C}b)^T W (\boldsymbol{Y} - \boldsymbol{X}\beta - \boldsymbol{C}b) + \lambda b^T \mathcal{D} b. \tag{6}$$

with $\boldsymbol{Y} = (R_{j1}, \ldots, R_{jT})^T$, $\boldsymbol{X} = (1, \boldsymbol{R}_M)$ and $\boldsymbol{R}_M = (R_{M1}, \ldots, R_{MT})$. Matrix C is a $n \times p$ dimensional spline basis built from rows

$$C_t = R_{Mt}(B_1(t), \ldots, B_p(t)), t = 1, \ldots, T$$

with $B_l(t)$ as l-th spline basis. Finally, $W = \text{diag}(w_{jt})$ is the diagonal matrix of weights accounting for heteroscedasticity and $\mathcal{D}$ is a penalty matrix steering with λ the smoothness of the resulting fit. Additional constraints on b are necessary to ensure identifiability, but for notational simplicity and since technically these do not cause problems we ignore them here (see Wood (2000) for a more technical insight). Keeping the smoothing parameter λ fixed, one gets an estimate for $\Theta = (\beta, b)$ via

$$\begin{aligned} \hat{\Theta} &= \left(\begin{pmatrix} \boldsymbol{X}^T \\ \boldsymbol{C}^T \end{pmatrix} \boldsymbol{W}(\boldsymbol{X}\boldsymbol{C}) + \lambda \text{diag}(0, \mathcal{D}) \right)^{-1} \begin{pmatrix} \boldsymbol{X}^T \\ \boldsymbol{C}^T \end{pmatrix} \boldsymbol{W}\boldsymbol{Y} \\ &=: \boldsymbol{M}_\lambda \boldsymbol{Y} \end{aligned}$$

Note that if weights w_{jt} depend on Θ as well, iterated estimation is necessary. In our fitting process we however confine ourselves to a two stage fitting routine only.

An important issue in non-parametric regression is the selection of an appropriate smoothing parameter. Choosing $\lambda = 0$ results in a wiggled estimate while $\lambda \to \infty$ would yield a time constant beta-coefficient. A widely accepted method for smoothing parameter selection (see Hastie and Tibshirani (1990)) is minimizing Akaike's information criterion (Akaike (1973)) or its relative the Generalized Cross Validation (Craven and Wahba (1979))

$$GCV(\lambda) = \frac{(\boldsymbol{Y} - \boldsymbol{X}\beta - \boldsymbol{C}b)^T \boldsymbol{W} (\boldsymbol{Y} - \boldsymbol{X}\beta - \boldsymbol{C}b) / T}{[1 - df(\lambda)/n]^2}$$

where $df(\lambda)$ is a measure for the degree of freedom or for the complexity of the fit, respectively. As motivated for instance in Hastie and Tibshirani (1990, ch. 3) a suitable choice for $df(\lambda)$ is the trace of the "hat" matrix, which here means

$$df(\lambda) = trace\{\boldsymbol{M}_\lambda(\boldsymbol{XC})\}$$

Clearly, $GCV(\lambda)$ does not allow for simple minimization, since $\partial GCV(\lambda)/\partial\lambda = 0$ does not provide an analytic solution. However, the simple idea of a Newton Raphson procedure can be applied to solve the first order derivative. Details are provided in Wood (2000) and the procedure is implemented in R.

When working with any automatic smoothing parameter selection one should be aware that the routines usually have a large variability. In practical terms this means, one should not blindly believe in the selected smoothing parameter and take the resulting fit for granted. Instead, one should visually check the fit with different smoothing parameters. If differences are minor, the automatic selected smoothing parameter can be accepted[17]. We followed this advice in our data example and found that the automatic selected smoothing parameters behaved satisfactory.

[17]See also Ruppert, Wand, and Carroll (2003b, ch. 5.4).

References

Akaike, H., 1973, "Information Theory and an Extension of the maximum likelihood principle," *Second international symposium on information theory*, pp. 267–281.

Alexander, G., and N. Chervany, 1980, "On the estimation and stability of beta," *Journal of Financial and Quantitative Analysis*, 15, 123–137.

Altman, E., B. Jacquillat, and M. Levasseur, 1974, "Comparative analysis of risk measures: France and the United States," *Journal of Finance*, 27, 1495–1511.

Baesel, J., 1971, "On the assessment of risk: Some further considerations," *Journal of Finance*, 27, 1491–1494.

Blume, M., 1971, "On the assessment of risk," *Journal of Finance*, 24, 1–10.

Bowman, A., and A. Azzalini, 1997, *Applied smoothing tecniques for data analysis* . Oxford Statistical Science Series, Clarendon Press, Oxford.

Campbell, J., and J. Cochrane, 2000, "Explaining the poor performance of consumption-based asset pricing models," *Journal of Finance*, 55, 1–16.

Cochrane, J., 2001, *Asset pricing*, Princeton University Press, Princeton.

Craven, P., and G. Wahba, 1979, "Smoothing noisy data with spline functions," *Numerical Mathematics*, 31, 377–403.

Dalgaard, P., 2002, *Introductory Statistics with R*, Springer Verlag.

De Boor, C., 1978, *A practical guide to Splines*, Springer Verlag, New York.

Deutsche-Börse-Group, 2003, *Leitfaden zu den Aktienindizes der Deutschen Börse*.

Green, P., and B. Silverman, 1994, *Nonparametric regression and generalized linear models* . Monographs on Statistics and Applied Probability, Chapman and Hall, London.

Grinblatt, M., and S. Titman, 2002, *Financial markets and corporate strategy*, McGraw-Hill/Irwin, New York, 2nd edn.

Gu, C., and G. Wahba, 1991, “Minimizing GCV/GML scores with multiple smoothing parameters via the newton method,” *SIAM Journal of Scientific and Statistical Computing*, 12, 383–398.

Hastie, T., 1996, “Pseudosplines,” *Journal of the Royal Statistical Society B*, 58, 379–396.

Hastie, T., and R. Tibshirani, 1990, *Generalized additive models* . Monographs on Statistics and Applied Probability, Chapman and Hall, London.

Hastie, T., and R. Tibshirani, 1993, “Varying-coefficient models,” *Journal of the Royal Statistical Society B*, 55, 757–796.

Kauermann, G., and J. Opsomer, 2004, “Generalized cross-validation for bandwidth selection of backfitting estimates in generalized additive models,” *Journal of Computational and Graphical Statistics*, 13, 66–89.

Markowitz, H., 1952, “Portfolio selection,” *The Journal of Finance*, 12, 77–91.

McCullagh, P., and J. Nelder, 1989, *Generalized linear models* . Monographs on Statistics and Applied Probability, Chapman and Hall, New York.

Opsomer, J., Y. Wang, and Y. Yang, 2001, “Nonparametric regression with correlated errors,” *Statistical Science*, 16, 134–153.

Roenfeldt, R., G. Griepentrong, and C. Pflaum, 1978, “Further evidence on the stationarity of beta coefficients,” *Journal of Financial and Quantitative Analysis*, 13, 117–121.

Ruppert, D., M. Wand, and R. Carroll, 2003a, *Semiparametric regression*, Cambridge University Press, Cambridge.

Ruppert, R., M. Wand, and R. Carroll, 2003b, *Semiparametric Regression*, Cambridge University Press.

Schaefer, S., R. Brealey, S. Hodges, and H. Thomas, 1975, “Alternative models of systematic risk,” In Elton, E. & M. Gruber, eds, *International capital markets*, North Holland, Amsterdam, 150–161.

Semmler, W., 2003, *Asset prices, booms and recessions: Financial markets, economic activity and the macroeconomy*, Springer Verlag, Berlin.

Sharpe, W., 1963, "A simplified model for portfolio analysis," *Management Science*, 9, 227–293.

Sharpe, W., 1964, "Capital asset prices: A theory of market equilibrium under conditions of risk," *Journal of Finance*, 19, 425–442.

Sharpe, W., 1970, *Portfolio theory and capital markets* . Series in Finance, McGraw-Hill, New York.

Theobald, M., 1981, "Beta stationarity and estimation period: Some analytical results," *Journal of Financial and Quantitative Analysis*, 16, 747–757.

Wells, C., 1995, *The Kalman filter in Finance*, Kluwer Academic Publishers, Dordrecht.

Wood, S., 2000, "Modelling and smoothing parameter selection with multiple quadratic penalties," *Journal of the Royal Statistical Society B*, 50, 413–436.

Wood, S., 2001, "mgcv: GAMs and generalized ridge regression for R," *R News*, 1/2, 20–25.