

Time-Varying Market Price of Risk in the CAPM – Approaches, Empirical Evidence and Implications

Christian M. Hafner

Helmut Herwartz *

January 1999

Abstract

Time-varying risk premia traditionally have been associated with the empirical fact that conditional second moments are time-varying. This paper additionally examines another possible source for time-varying risk premia, namely the market price of risk (λ). For utility functions that do not imply constant risk aversion measures, the market price of risk will in general change over time. We provide empirical evidence for the German stock market in a bivariate GARCH-M framework using alternative specifications for λ . The results indicate that a model with λ being a function of typical volatility measures performs best for most series. To facilitate the interpretation of the results, we plot impulse response functions of the risk premia.

Keywords: Market price of risk, Multivariate GARCH-Models, impulse response analysis, CAPM

*Institut für Statistik und Ökonometrie and Sonderforschungsbereich 373, Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin, Spandauer Str. 1, D-10178 Berlin, Germany. e-mail hafner@wiwi.hu-berlin.de or helmut@wiwi.hu-berlin.de. Financial support by the Deutsche Forschungsgemeinschaft and helpful comments of an anonymous referee are gratefully acknowledged.

1 Introduction

For at least a decade now, there have been no doubts about the empirical evidence for time-varying risk premia of financial assets. So far, this stylized fact was mainly attributed to the time-varying behavior of conditional second moments. For example, Engle, Lilien and Robins (1987) establish a link between the risk premium and the ARCH-type volatility, the so-called ARCH-M model. For the capital asset pricing model (CAPM), Bollerslev, Engle and Wooldridge (1988) introduce time-varying covariances to obtain time-varying betas and thus time-varying risk premia using multivariate generalized ARCH (GARCH) models. The increasing experience with multivariate GARCH models over past years has led to more adequate volatility specifications (Hafner and Herwartz, 1998a). However, to obtain a feasible econometric model in the CAPM, a typical assumption is that the market price of risk, the so-called *lambda*, is constant over time. In this paper, we argue that this assumption may be too restrictive. There are two scenarios for which the market price of risk is time-varying: First, for utility functions that imply both absolute and relative risk aversion to be dependent on the return, market price of risk is in general a function of the conditional first and second moment of the return. Second, if the utility function has parameters that are time-varying and determine the degree of risk aversion. We provide examples for both scenarios.

We give empirical evidence of time-varying market price of risk for the German stock market. To this end, we use a multivariate GARCH framework as in Bollerslev, Engle, and Wooldridge (1988) and Hafner and Herwartz (1998a). For *lambda*, we employ various specifications. The result of the empirical part is that for the majority of analyzed series a model for which *lambda* depends on lagged squared innovations outperforms models with constant *lambda*. This suggests that there is a link between the market price of risk and typical volatility measures. The parameter estimates imply a positive relation between *lambda* and lagged squared innovations. The interpretation is that for large lagged innovations, not only volatility increases, but also *lambda*. Thus, there is a double effect on the risk premium, the one stemming from volatility, the other from a time-varying *lambda*.

In order to analyze the consequences for the risk premium, we further suggest an impulse response methodology as in Hafner and Herwartz (1998b). For the preferred specification of *lambda*, the risk premium is a simple linear function of volatility and lagged squared innovation. Thus, impulse response analysis can be performed by computing conditional expectations of the risk premium.

2 The CAPM with Time-varying Market Price of Risk

In the standard CAPM framework with riskfree rate r_f , n risky assets with price vector $S_t = (S_{1,t}, \dots, S_{n,t})'$ and the vector of gross returns $r_t = (r_{1,t}, \dots, r_{n,t})'$, with $r_{i,t} = S_{i,t}/S_{i,t-1}$, the basic equilibrium equation is

$$E_{t-1}[r_t - r_{f,t}] = \lambda \Sigma_t w_{t-1}. \quad (1)$$

In (1), Σ_t is the covariance matrix of the risky assets at time t conditional on the information set available at time $t - 1$, Ψ_{t-1} , and w_t is the weight vector of the assets in the market portfolio at time t . The parameter λ is the aggregated coefficient of risk aversion and is sometimes referred to as lambda, the market price of risk. For the assumptions underlying the CAPM, the individual expected utilities are functions only of the mean and variance of the returns. Thus, the expected utility $E[U(r_{m,t})]$ of the representative agent can be denoted by $v(\mu_t, \sigma_{m,t}^2)$ and in equilibrium

$$\lambda = -2 \frac{\partial v / \partial \sigma_{m,t}^2}{\partial v / \partial \mu_t}.$$

By denoting $\mu_t = E_{t-1}[r'_t w_{t-1}]$ and $\sigma_{m,t}^2 = w'_{t-1} \Sigma_t w_{t-1}$ the conditional mean and conditional variance of the market portfolio, respectively, we can specifically write from (1) for the market portfolio

$$E_{t-1}[r_{m,t} - r_{f,t}] = \lambda \sigma_{m,t}^2 \quad (2)$$

and by substituting for λ in (1)

$$E_{t-1}[r_t - r_{f,t}] = \beta_t E_{t-1}[r_{m,t} - r_{f,t}] \quad (3)$$

with the vector $\beta_t = \Sigma_t w_{t-1} / \sigma_{m,t}^2$. This is the well known market beta form of the CAPM, which is due to Sharpe (1964) and Lintner (1965). It remains valid when we assume λ to be a time-varying function of the past, λ_t , because it obviously cancels out in the derivation of (3), see also Gouriéroux (1997, p. 187). However, the correct assumption concerning λ remains an important issue when estimating (3), because the expected market return is not observed. In the literature, one traditionally assumed that λ is constant. It is well known that this can be justified in the following cases:

1. The representative agent has constant relative risk aversion and logarithmic returns are normally distributed.
2. The representative agent has constant absolute risk aversion and gross returns are normally distributed.

One can imagine, however, that in more general situations the risk aversion parameter λ is not constant but a function of the past. We give two examples of such situations:

1. The utility function of the representative agent is of the form

$$U(r_{m,t}) = \gamma r_{m,t} - \exp(-\theta r_{m,t}).$$

For $\gamma = 0$, we obtain the special case of exponential utility, i.e. constant absolute risk aversion, and for $\theta = 0$ the linear utility, i.e. risk neutrality. For the case of conditionally normally distributed returns, $N(\mu_t, \sigma_{m,t}^2)$, we obtain

$$E[U(r_{m,t})] = \gamma \mu_t - \exp(-\theta \mu_t + \theta^2/2\sigma_{m,t}^2)$$

and

$$\lambda_t = 2 \frac{\partial v / \partial \sigma_{m,t}^2}{\partial v / \partial \mu_t} = \frac{\theta^2}{\gamma \exp(\theta \mu_t - \theta^2/2\sigma_{m,t}^2) + \theta}.$$

Here, λ_t is a function of the past through the conditional moments μ_t and $\sigma_{m,t}^2$.

2. The utility function of the representative agent is of the power form, say, with time dependent relative risk aversion a_t :

$$U_t(r_{m,t}) = \frac{r_{m,t}^{1-a_t} - 1}{1 - a_t}$$

where a_t is a function of the past. This can be motivated, for example, by habit persistence. For $a_t \rightarrow 0$, logarithmic utility is obtained as a special case.

In general, the aggregated risk aversion parameter λ_t will therefore be time-varying and this should be taken into account when specifying the econometric model of the CAPM.

An econometric specification of the CAPM augmented by an intercept term is given for asset i as

$$r_{i,t} = \gamma + E_{t-1}[r_{f,t}] + \beta_{i,t} E_{t-1}[r_{m,t} - r_{f,t}] + \varepsilon_{i,t} \quad (4)$$

with the asset's 'beta' $\beta_{i,t} = \sigma_{im,t} / \sigma_{m,t}^2$ measuring the undiversifiable risk associated with a specific asset. The inclusion of γ in (4) does not follow from the CAPM, hence, standard specification tests of the CAPM amount to test against whiteness of estimated residuals and against $H_0 : \gamma = 0$. Note that in the form (4) the model cannot be estimated because both conditional expectations on the right hand side are unobserved. For the riskfree rate we can assume $E_{t-1}[r_{f,t}] = r_{f,t-1}$, but for $E_{t-1}[r_{m,t} - r_{f,t}]$ we have to resort to the form (2) by specifying λ and $\sigma_{m,t}^2$.

Time varying variances and covariances may be introduced assuming the bivariate error sequence $\varepsilon_t = (\varepsilon_{i,t}, \varepsilon_{m,t})'$ to exhibit conditional heteroskedasticity, i.e.

$$\Sigma_t = E[\varepsilon_t \varepsilon_t' | \Psi_{t-1}] = \begin{pmatrix} \sigma_{i,t}^2 & \sigma_{im,t} \\ \sigma_{im,t} & \sigma_{m,t}^2 \end{pmatrix}.$$

As a parametric specification of Σ_t one may adopt a multivariate GARCH-type model. This turned out to be useful in many previous empirical studies of multivariate financial time series.

With the above considerations, the market-price of risk given in (2) can be written as

$$\lambda_t = \frac{E_{t-1}[r_{m,t}] - r_{f,t-1}}{\sigma_{m,t}^2} \quad (5)$$

and is assumed here to be time-varying, as motivated above. This contrasts previous specifications of λ_t being constant over time as in Bollerslev, Engle and Wooldridge (1988) and Hafner and Herwartz (1998a). We will suggest some specifications for λ_t in Section 4.

Defining the bivariate excess return series $y_t = (r_{i,t} - r_{f,t-1}, r_{m,t} - r_{f,t-1})'$ one obtains the bivariate model

$$y_t = \gamma + \lambda_t \begin{pmatrix} \sigma_{im,t} \\ \sigma_{m,t}^2 \end{pmatrix} + \varepsilon_t, \quad (6)$$

which can be interpreted as a GARCH-M model (Engle, Lilien, Robins, 1987) with time-varying coefficient.

Instead of estimating the CAPM by means of a set of bivariate equations one may regard a system of “seemingly unrelated” equations collecting $r_{i,t} - r_{f,t-1}, i = 1, \dots, n$ and $r_{m,t} - r_{f,t-1}$ as a competing econometric device. From such a system representation a unique estimate of λ_t could be obtained. With respect to estimation efficiency, however, a simultaneous estimation is not expected to improve the set of equations form in (6). Note that the error terms in ε_t are supposed to capture unsystematic risk which should hardly show any correlation pattern across assets and thus we refrained from performing a system estimation (see e.g. Judge et al., 1988, Chapter 11).

Due to the complicated iterative procedure necessary to estimate (6) one may also regard the system estimation of the CAPM with time varying coefficients to be unfeasible in practice. To address the issue of estimating the market’s λ_t we provide a brief illustration of estimated λ_t processes stemming from investigations of different assets using (6) within the discussion of our estimation results.

Candidate parametric models for Σ_t and (quasi-)maximum-likelihood estimation (QML) of the model in (6) will be outlined in the next section.

3 Bivariate GARCH-type Models – Specification and Estimation

The generalization of univariate (G)ARCH-type models of conditional heteroskedasticity (see Engle, 1982, and Bollerslev, 1986) to the bivariate case is more or less straightforward. The two-dimensional random vector $\varepsilon_t = (\varepsilon_{i,t}, \varepsilon_{m,t})'$ may be written as

$$\varepsilon_t = \Sigma_t^{1/2} \xi_t \quad (7)$$

where ξ_t denotes an *i.i.d.* random vector with mean zero and covariance matrix equal to the bivariate identity matrix (I_2). Conditional on Ψ_{t-1} , the elements of Σ_t are completely determined by their own history Σ_{t-i} , $i = 1, \dots, p$, and lagged observations ε_{t-i} , $i = 1, \dots, q$. The so-called vec-specification of the multivariate GARCH(p, q) model provides the dynamics of the elements of the lower fraction of Σ_t , i.e. $\text{vech}(\Sigma_t)$:

$$\text{vech}(\Sigma_t) = c + \sum_{i=1}^q A_i \text{vech}(\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{i=1}^p G_i \text{vech}(\Sigma_{t-i}). \quad (8)$$

In (8), A_i and G_i denote (3×3) matrices. Additionally, 3 parameters in the vector c account for time invariant variance components. Since QML estimation of GARCH-type models involves non-linear optimization routines one may imagine that even for the multivariate GARCH(1,1) model the vec-specification easily becomes intractable. The dimension of the relevant parameter space may be reduced e.g. by assuming the matrices A_i and G_i to be diagonal as adopted e.g. by Bollerslev, Engle, and Wooldridge (1988) such that the (k, l) -element in Σ_t depends linearly on the respective elements of the matrices $\varepsilon_{t-i} \varepsilon'_{t-i}$ and Σ_{t-i} . However, the diagonal vec-model *a-priori* excludes possibly important cross dynamics relating one variable's conditional volatility on lagged innovations observed for another variable. A more general structure allowing for interdependence is given by the so-called BEKK-model (Baba, Engle, Kraft and Kroner, 1990),

$$\Sigma_t = C'_0 C_0 + \sum_{k=1}^K \sum_{i=1}^q A'_{ki} \varepsilon_{t-i} \varepsilon'_{t-i} A_{ki} + \sum_{k=1}^K \sum_{i=1}^p G'_{ki} \Sigma_{t-i} G_{ki}, \quad (9)$$

where C_0 is an upper triangular matrix and A_{ki} and G_{ki} are 2×2 parameter matrices. Even in the case $K = 1$, the model in (9) relates each element of Σ_t to all elements in $\varepsilon_{t-i} \varepsilon'_{t-i}$ and Σ_{t-i} . Note that (9) ensures Σ_t to be positive definite without imposing further parameter restrictions. Engle and Kroner (1995) discuss the BEKK-model in detail. For the present analysis we take $K = 1$ and concentrate on the GARCH(1,1) model. In this case the assumption that the upper left elements of A_{11} and G_{11} are greater than zero is sufficient for the model parameters to be identified. As in Hafner and Herwartz (1998a) we adopt extensions of the symmetric GARCH-models given above in order to allow the potential of a larger impact of bad news (negative lagged innovations) compared with good news (positive lagged innovations) on volatility. This empirical phenomenon is known since Black (1976) and is frequently called 'leverage effect'. It may be regarded as a stylized fact of conditional variances of risky assets. A comprehensive list of the variance specifications under study reads as follows:

$$\text{M1: } \Sigma_t = C'_0 C_0 + A'_{11} \varepsilon_{t-1} \varepsilon'_{t-1} A_{11} + G'_{11} \Sigma_{t-1} G_{11} \quad (10)$$

$$\text{M2: } \Sigma_t = C'_0 C_0 + A'_{11} \varepsilon_{t-1} \varepsilon'_{t-1} A_{11} + A'_{21} \varepsilon_{t-1} \varepsilon'_{t-1} A_{21} I_{\varepsilon_{1,t-1} < 0} + G'_{11} \Sigma_{t-1} G_{11} \quad (11)$$

$$\text{M3: } \Sigma_t = C'_0 C_0 + A'_{11} \varepsilon_{t-1} \varepsilon'_{t-1} A_{11} + A'_{31} \varepsilon_{t-1} \varepsilon'_{t-1} A_{31} I_{\varepsilon_{2,t-1} < 0} + G'_{11} \Sigma_{t-1} G_{11} \quad (12)$$

$$\begin{aligned} \text{M4: } \Sigma_t &= C'_0 C_0 + A'_{11} \varepsilon_{t-1} \varepsilon'_{t-1} A_{11} \\ &+ A'_{21} \varepsilon_{t-1} \varepsilon'_{t-1} A_{21} I_{\varepsilon_{1,t-1} < 0} + A'_{31} \varepsilon_{t-1} \varepsilon'_{t-1} A_{31} I_{\varepsilon_{2,t-1} < 0} + G'_{11} \Sigma_{t-1} G_{11} \end{aligned} \quad (13)$$

In (11) to (13), $I_{(\cdot)}$ denotes the indicator function. Note that these models may be regarded as natural extensions of the univariate threshold GARCH-model introduced by Glosten, Jagannathan and Runkle (1993) and Zakoian (1994). However, one may think of other devices to take asymmetry into account. For example, Braun, Nelson and Sunier (1995) introduce the bivariate exponential GARCH-model. Engle and Ng (1993) provide an empirical comparison of the GARCH and exponential GARCH model in the univariate case. They conclude that empirically both the threshold and the exponential GARCH applied to a Japanese stock index series perform similarly, that the EGARCH model however tends to overweigh the impact of large innovations on volatility, due to the exponential increase of the news impact curve. Also, an impulse response analysis of volatility is easier to do for the additive TGARCH than for the multiplicative EGARCH model. Since news might occur in the system through each of the components of ε_t and thus of ξ_t simultaneously or separately, we distinguish models M2 to M4 as asymmetric counterparts of the symmetric specification M1.

Engle and Kroner (1995) state that for each BEKK model there is a unique equivalent vec-representation. Thus, when discussing the properties of M1 to M4, we can also consider the equivalent vec-specification by defining $c^* = (C_0 \otimes C_0)' \text{vec}(I_2)$, $A_1^* = (A_{11} \otimes A_{11})'$, $A_2^* = (A_{21} \otimes A_{21})'$, $A_3^* = (A_{31} \otimes A_{31})'$, and $G_1^* = (G_{11} \otimes G_{11})'$. After eliminating from c^* , A_1^* , A_2^* , A_3^* , and G_1^* those rows and columns that are superfluous due to the symmetry of covariances, one obtains the matrices c , A_1 , A_2 , A_3 , and G_1 as in (8). This transformation is notationally more convenient and is consistent with the next section where we discuss impulse response functions only for the vec-specification. The following proposition provides a result for the covariance stationarity of the model M4. The result applies immediately to M1, M2 and M3 by setting the corresponding matrices A_2 and/or A_3 to zero.

Proposition 1 *Assume that both components of the error ξ_t have a symmetric density around zero. Then the process ε_t is covariance stationary if and only if all eigenvalues of*

$$A_1 + A_2/2 + A_3/2 + G_1 \tag{14}$$

have moduli less than one. For the implied unconditional covariance matrix Σ , one obtains

$$\text{vech}(\Sigma) = E[\text{vech}(\varepsilon_t \varepsilon_t')] = (I_3 - A_1 - A_2/2 - A_3/2 - G_1)^{-1} c. \tag{15}$$

Proof: Hafner and Herwartz (1998a).

The elements of the parameter matrices in (10) to (13) and the additional parameters in (6) are conveniently estimated by numerical procedures. Within this study we used the BHHH-algorithm as described e.g. in Judge et al. (1988) to maximize the quasi log-likelihood function derived under the assumption of normally distributed innovations

ξ_t . Under normality, the contribution of an observation

$$\varepsilon_t = y_t - \gamma - \lambda_t \begin{pmatrix} \sigma_{im,t} \\ \sigma_{m,t}^2 \end{pmatrix}$$

to the joint log-likelihood of a sample with T observations ($\log L = \sum_{t=1}^T l_t$) is obtained as:

$$l_t = -\ln(2\pi) - \frac{1}{2} \ln |\Sigma_t| - \frac{1}{2} \varepsilon_t' \Sigma_t^{-1} \varepsilon_t. \quad (16)$$

Although consistency of the QML-estimation has not been proven yet for the multivariate case we conjecture such a result along the lines given in Bollerslev and Wooldridge (1992) and Lumsdaine (1996) for the univariate case.

4 Specifications of the Market Price of Risk and Empirical Results

We investigate daily prices of 21 German stocks for the period January 2, 1990 to December 30, 1996 including 1752 trading days. Stock price data were obtained from the *Deutsche Kapitalmarktdatenbank*, Karlsruhe. Returns on the market portfolio were computed using the so-called DAFOX index series which is provided by the University of Karlsruhe. This index is computed for research purposes and is composed of all stocks traded at the Frankfurt stock exchange. All stock market series were adjusted for payments out of the stock and for changes of their nominal value.

A money market interest rate for deposits with one month time to maturity was chosen to approximate risk free returns. Daily rates were provided by the Deutsche Bundesbank. The interest rate and the DAFOX series are given in Figure 1. Our sample covers a period of a relatively high interest rate indicating the huge demand for liquidity in the sequel of the German unification. Thus a period of negative excess returns may be conjectured for the beginning of our sample. The second half of our sample period is characterized by a marked upward trending evolution of stock prices as it was observed for most major stock markets.

Adopting a univariate analysis, Hafner and Herwartz (1999) show for the same data set that additional to time-varying risk premia an autoregressive component is often helpful to explain the degree of autocorrelation. For this reason, we augment the bivariate GARCH-M model in (6) by a 2×2 Matrix B capturing autoregressive dynamics of the observed excess rates of return:

$$y_t = \gamma + B y_{t-1} + \lambda_t \begin{pmatrix} \sigma_{im,t} \\ \sigma_{m,t}^2 \end{pmatrix} + \varepsilon_t. \quad (17)$$

An essential step for estimating the GARCH-M model is the specification of an appropriate bivariate volatility model. To select one of the alternative specifications (M1 to M4)

listed in Section 3 we estimated the GARCH-M model in (17) assuming λ to be constant with all competing volatility models. Numerical procedures written in GAUSS 3.2 were used to perform QML estimation of (17). The resulting values of the log-likelihood function are provided in Table 1. Note that the most general model (M4) has 4 (8) additional parameters relative to M2 and M3 (M1). Without relying too much on formal tests we selected M4 as a convenient volatility model for a given series in question if the value of its log-likelihood exceeded the respective measure of M2 and M3 by at least 5.0 points. The symmetric model (M1) is clearly rejected for almost all series under study relative to the remaining specifications. Since M2 and M3 are comprised by the same number of model parameters, the choice between these models was determined by comparing the respective maximum values of the log-likelihood function. In Table 1 selected variance specifications are indicated with an asterisk. To provide some insight into the relevance of autocorrelation for the series under study, the first column of Table 1 provides the values of the log-likelihood obtained if the matrix B in (17) is restricted to be a zero matrix. The log-likelihood values of the restricted model may be directly compared with the selected specification of a GARCH-M model including B . Neglecting autocorrelation involves a loss measured by means of the log-likelihood which is significant at all conventional levels for almost all series under study.

To allow for time dependence of λ_t we adopted the following parametric specifications:

$$\begin{aligned}
\text{L1: } \lambda_t &= \alpha_0 + \alpha_1 \sigma_{m,t}^2 \\
\text{L2: } \lambda_t &= \alpha_0 + \alpha_1 \varepsilon_{m,t-1}^2 / \sigma_{m,t}^2 \\
\text{L3: } \lambda_t &= \alpha_0 + \alpha_1 \varepsilon_{m,t-1}^2 / \sigma_{m,t-1}^2 \\
\text{L4: } \lambda_t &= \alpha_0 + \alpha_1 |\varepsilon_{m,t-1}| \\
\text{L5: } \lambda_t &= \alpha_0 + \alpha_1 \sigma_{m,t}^2 I_{(\varepsilon_{m,t-1} < 0)} \\
\text{L6: } \lambda_t &= \alpha_0 + \alpha_1 \varepsilon_{m,t-1}^2 / \sigma_{m,t}^2 I_{(\varepsilon_{m,t-1} < 0)} \\
\text{L7: } \lambda_t &= \alpha_0 + \alpha_1 \varepsilon_{m,t-1}^2 / \sigma_{m,t-1}^2 I_{(\varepsilon_{m,t-1} < 0)} \\
\text{L8: } \lambda_t &= \alpha_0 + \alpha_1 |\varepsilon_{m,t-1}| I_{(\varepsilon_{m,t-1} < 0)} \\
\text{L9: } \lambda_t &= \alpha_0 + \alpha_1 \sigma_{m,t}^2 I_{(\varepsilon_{m,t-1} < 0)} + \alpha_2 \sigma_{m,t}^2 I_{(\varepsilon_{m,t-1} > 0)} \\
\text{L10: } \lambda_t &= \alpha_0 + \alpha_1 \varepsilon_{m,t-1}^2 / \sigma_{m,t}^2 I_{(\varepsilon_{m,t-1} < 0)} + \alpha_2 \varepsilon_{m,t-1}^2 / \sigma_{m,t}^2 I_{(\varepsilon_{m,t-1} > 0)} \\
\text{L11: } \lambda_t &= \alpha_0 + \alpha_1 \varepsilon_{m,t-1}^2 / \sigma_{m,t-1}^2 I_{(\varepsilon_{m,t-1} < 0)} + \alpha_2 \varepsilon_{m,t-1}^2 / \sigma_{m,t-1}^2 I_{(\varepsilon_{m,t-1} > 0)} \\
\text{L12: } \lambda_t &= \alpha_0 + \alpha_1 |\varepsilon_{m,t-1}| I_{(\varepsilon_{m,t-1} < 0)} + \alpha_2 |\varepsilon_{m,t-1}| I_{(\varepsilon_{m,t-1} > 0)}
\end{aligned}$$

All specifications under study relate λ_t to the history of the return process, Ψ_{t-1} . For none of the models we imposed a positivity constraint, so negative estimates of λ_t may be interpreted as evidence against the CAPM. We will come back to this issue in Section 5. The model L1 (L4) states λ_t to be a linear function in the conditional variance (absolute lagged return) of the market portfolio. In model L3, λ_t is related to lagged

squared innovations, $\varepsilon_{m,t-1}^2/\sigma_{m,t-1}^2$. Under conditional normality of ε_t and assuming the employed volatility process to represent the true second order moments squared innovations are i.i.d. and conditionally follow a $\chi^2(1)$ distribution. However, with respect to computational feasibility L3 turned out to suffer from numerical difficulties. Extremely large values of λ_t are obtained for this specification if large values of $\varepsilon_{m,t-1}$ occur in states of the dynamic system in which their conditional variance is relatively low. To cope with numerical problems, L2 may be regarded as a close approximation to L3 for almost all observations in the sample. Note that the quantity $\varepsilon_{m,t-1}^2/\sigma_{m,t}^2$ should always be conveniently bounded, since its denominator is computed partly from its numerator. Closely related to L1 to L4 are the specifications L5 to L8 which propose the linear relations suggested above to hold only for those states of the system where lagged observed innovations are negative. In case of ‘good news’ hitting the stock market λ_t is assumed to be constant. Different slope coefficients for the linear relationships given in L1 to L4 with respect to ‘good news’ and ‘bad news’ occurring in $t - 1$ are allowed within the representations L9 to L12. Of course, to make estimation of the competing devices feasible the right hand side variables in L1 to L12 have to be replaced by their estimates conditional on Ψ_{t-1} .

In Table 2 diagnostic results for models with time varying lambda are provided. We report twice the difference of the log-likelihood of the estimated specifications L1 to L12 relative to the CAPM with lambda assumed to be constant ($\lambda_t = \lambda$). Notice that the specifications L1 to L8 have one additional parameter relative to the restricted model. A further parameter is introduced in L9 to L12. Although we view our QML diagnostic results as more or less descriptive, entries which are larger than 4 (6) are indicated in Table 2 with an asterisk for estimated specifications L1 to L8 (L9 to L12). Note that these ‘critical values’ would roughly correspond to a 5% significance level if the statistics were regarded as formal tests. In principle, all entries in Table 2 should be positive. Small but negative statistics are due to numerical problems involved with the maximization of the log-likelihood function in a very large parameter space. For 13 of 21 series under study promising improvements of the restricted model are obtained if lambda is allowed to depend on the history of the bivariate process. Simply by counting ‘significant’ statistics obtained within related specifications it turns out that 11, 16 and 10 improvements are obtained for L1 to L4, L5 to L8, and L9 to L12, respectively. This result supports the case for asymmetry of the dependence of λ_t . L5 to L6 have in common that the stated linear relationship for λ_t holds only if bad news hit the market at time $t - 1$. Within the specifications L5 to L8 it turns out that in most cases (10 of 16) L6 provides considerable improvements of the standard specification with lambda being constant through time. L6 relates λ_t to squared innovations $\varepsilon_{m,t-1}^2/\sigma_{m,t}^2$. Closely related to this specification are L2 and L10 which also perform considerably well comparing results obtained for L1 to L4 and L9 to L12 respectively. As mentioned above these models may be regarded as an approximation to specifications explaining λ_t by means of estimated squared innovations

$\varepsilon_{m,t-1}^2/\sigma_{m,t-1}^2$. Assuming λ_t to be linear in $\sigma_{m,t-1}^2$ or $|\varepsilon_{m,t-1}|$ amounts to minor improvements of the restricted model relative to the assumption of linearity in lagged squared innovations.

5 Impulse Response Analysis of the Risk Premium

For the models suggested in the previous section for the market price of risk, we can now proceed to investigate the impact of independent innovations on the risk premium. In general, we will distinguish two different sources of innovations: asset specific and market innovations. They are represented by the stochastically independent innovations $\xi_{i,t}$ and $\xi_{m,t}$ in model (7). Economically, this independence can be justified if the weight of each asset in the market portfolio is negligibly small. Recall that our series that represents the market portfolio, the DAFOX, covers all traded assets at the Frankfurt stock exchange, so that a potential dependence of asset specific and market innovations is reduced as much as possible.

In our general framework, the risk premium consists of two time varying components: the volatility part and the price of risk part. It is thus not *ex ante* clear how the product of both components reacts to positive or negative innovations. In fact, it may be that volatility increases for large innovations (as is the case in our GARCH framework) but that the price of risk decreases. In this case it depends on the magnitude of both effects to evaluate whether the risk premium increases or decreases. On the other hand, it may be that for large innovations also the price of risk increases, which would imply an even stronger increase of the risk premium.

A particularly simple form for the risk premium is obtained for the model (L2),

$$\lambda_t = \alpha_0 + \alpha_1 \varepsilon_{m,t-1}^2 / \sigma_{m,t}^2,$$

since then we have for the risk premium

$$p_t := E_{t-1}[r_{m,t}] - r_{f,t-1} = \lambda_t \sigma_{m,t}^2 = \alpha_0 \sigma_{m,t}^2 + \alpha_1 \varepsilon_{m,t-1}^2. \quad (18)$$

We define the impulse response function for the risk premium as

$$P_k(\xi_t) = E[p_{t+k} \mid \xi_t, \Sigma_t] \quad (19)$$

for $k = 1, 2, \dots$ and $\xi_t = (\xi_{i,t}, \xi_{m,t})'$ the independent innovations to the system as given in (7). Thanks to the independence of the components of ξ_t , one may consider arbitrary shock scenarios. An alternative, due to Gallant, Rossi and Tauchen (1993), lets shocks occur in the conditionally dependent ε_t . Since GARCH models are linear in ε_t^2 , a related approach considers the derivatives of volatility forecasts with respect to squared ε_t , as in Baillie, Bollerslev and Mikkelsen (1996) for univariate ARCH(∞)-type processes.

In a multivariate framework, however, considering shocks in ε_t involves the additional task to determine realistic scenarios that take into account the contemporaneous correlation of the variables. It may be useful to emphasize that we do not have this problem for our impulse response function. For instance, one innovation may be restricted to zero to analyze a nonzero innovation in the other component.

Koop, Pesaran and Potter (1996) use Monte Carlo techniques to generate the distribution of impulse responses conditional on initial conditions, an initial shock, intermediate innovations and the model parameters, all of which can be regarded as random variables. This approach may provide valuable structural insights into the process dynamics if the conditional moments can not be determined analytically.

Unlike Gallant, Rossi and Tauchen (1993) and Koop, Pesaran and Potter (1996), we do not include a baseline function in (19), so our impulse response function for the risk premium will approach the unconditional risk premium rather than zero, provided that p_t is stationary.

For the risk premium in (18), we obtain

$$\begin{aligned} P_k(\xi_t) &= \alpha_0 E[\sigma_{m,t+k}^2 \mid \xi_t, \Sigma_t] + \alpha_1 E[\varepsilon_{m,t+k-1}^2 \mid \xi_t, \Sigma_t] \\ &= \alpha_0 E[\sigma_{m,t+k}^2 \mid \xi_t, \Sigma_t] + \alpha_1 E[\sigma_{m,t+k-1}^2 \mid \xi_t, \Sigma_t] \\ &= \alpha_0 V_{m,k}(\xi_t) + \alpha_1 V_{m,k-1}(\xi_t) \end{aligned} \tag{20}$$

where $V_k(\xi_t) = E[\sigma_{m,t+k}^2 \mid \xi_t, \Sigma_t]$ denotes the volatility impulse response function as introduced by Hafner and Herwartz (1998b). As condition on Σ_t we consider the steady state, i.e. $\Sigma_t = \Sigma$. This is not a crucial restriction, because varying the state of Σ_t only affects the level of P_k interpreted as a function of ξ_t , but not its typical shape. We see that the impulse response function for the risk premium for this particular model (L2) is just a linear combination of volatility impulse response functions. These are nonlinear functions of ξ_t , but they can be calculated analytically. For example, for the vec-GARCH(1,1) model, we have

$$V_1(\xi_t) = c + A_1 \text{vech}(\Sigma^{1/2} \xi_t \xi_t' \Sigma^{1/2}) + G_1 \text{vech}(\Sigma),$$

and, for $k \geq 2$,

$$V_k(\xi_t) = c + (A_1 + G_1) V_{k-1}(\xi_t).$$

In the limit, $P_k(\xi_t)$ approaches the unconditional risk premium, which is for model L2 the rescaled unconditional market variance σ_m^2 :

$$\lim_{k \rightarrow \infty} P_k(\xi_t) = (\alpha_0 + \alpha_1) \sigma_m^2.$$

For the case $\alpha_0 + \alpha_1 = 0$, the impulse response for the risk premium converges to zero. This case could be interpreted as unconditional risk neutrality, whereas conditionally the representative agent may still reveal risk aversion or risk loving behavior, depending on the sign of α_0 and α_1 , and on the evolution of $V_{m,k}(\xi_t)$.

When we are interested in impulse response functions for the threshold models (L6) and (L10), we have to make an assumption concerning the symmetry of the distribution of ξ_t . For the symmetric case, we obtain for (L6)

$$P_k(\xi_t) = \alpha_0 V_{m,k}(\xi_t) + \frac{\alpha_1}{2} V_{m,k-1}(\xi_t),$$

and for (L10)

$$P_k(\xi_t) = \alpha_0 V_{m,k}(\xi_t) + \frac{\alpha_1 + \alpha_2}{2} V_{m,k-1}(\xi_t).$$

The estimated impulse response functions for the bivariate series ALLIANZ-DAFOX and DAIMLER-DAFOX are given in Figure 5. There are two independent innovations in the vector ξ_t and we choose an isolated point of view by restricting one at time t to be zero, the other to vary. The left axes show in the panels on the left $\xi_{i,t}$, i.e. an innovation to the asset, in the panels on the right they represent an innovation to the market, $\xi_{m,t}$. The functions are plotted for fifty time periods.

First, notice that the unconditional lambda for ALLIANZ is negative, which in the light of the CAPM appears very unusual, because it would imply risk loving behavior. In fact, this may even be viewed as an inconsistency with the standard CAPM model. Negative lambdas were found for the majority of analyzed series. However, recall from Figure 1 that the beginning of the time period, 1990 until 1993, was characterized by the effects of the German unification, rising interest rates and stagnating stock prices. For the end of the sample period, lambdas are predominantly positive, so one should consider longer samples to infer against the CAPM.

For ALLIANZ, both plots show a similar pattern: the risk premium tends to decrease when the innovations are negative, while it remains at about the same level for positive innovations. This asymmetry arises from the threshold GARCH specification for volatility, because for ALLIANZ we have chosen the double asymmetric specification M4. For the chosen model L2, the parameter estimates are such that α_0 is negative with larger absolute value than α_1 , which is positive. Since there is high persistence in volatility (the eigenvalues of the matrix $A_1 + A_2/2 + A_3/2 + G_1$ in (14) are close to one) volatility is slowly changing over time and we see from (20) that the function will behave similarly to the corresponding volatility impulse response, but with negative sign.

For DAIMLER (lower plots), the unconditional lambda is positive and the risk premium impulse responses thus display a pattern similar to volatility impulse responses. The function for the asset-specific innovation increases at the first periods, since we restricted the second component in ξ_t (i.e. the market innovation) to be zero, thus volatility is underestimated. The variation in the response functions to asset-specific news is less than the variation caused by market-specific news, but the persistence is higher. Note also that the preferred volatility specification was double-asymmetric (M4) for ALLIANZ and market-innovation asymmetric (M3) for DAIMLER. This is the reason for the ALLIANZ functions to be both asymmetric, whereas for DAIMLER only the response to market

innovations is asymmetric. The asymmetry is caused by the volatility leverage effect. Obviously there are inverse effects for the risk premium depending on the lambdas: when lambdas are positive, the risk premium behaves similar to volatility, so there is the usual leverage effect also for the risk premium. When lambda is negative, signs revert and risk premia decrease strongly for negative innovations.

In the light of the CAPM, we would expect to obtain estimated lambdas that are the same, or at least very similar, for all series. To give an example, we plotted the estimated lambda series for BASF, PREUSSAG (Figure 4) and ALLIANZ (Figure 3). For BASF and PREUSSAG, the lambdas look very similar, but for ALLIANZ it somewhat differs. In particular, the large peaks occur at different times: For BASF/PREUSSAG the largest peak goes along with a large increase in volatility (upper panel of Figure 3), whereas for ALLIANZ the largest peak occurs in a low-volatility state. This may also be explained by the inverse relation of risk premia and volatility for ALLIANZ.

6 Conclusions

We have generalized the standard empirical methodology of estimating the CAPM with time-varying covariances to allow also for time-varying market price of risk (lambda). We tried several alternative specifications for lambda and found significant improvement of the likelihood results for the majority of the analyzed German stock returns. Surprisingly, we found a negative unconditional lambda *inter alia* for ALLIANZ, which implies an inverse relation of volatility and risk premium. This may be viewed as an inconsistency with the assumptions of the CAPM. Impulse response functions for risk premia show that the primary shape of these functions is determined by the volatility specification with sign according to the sign of the unconditional lambda. For negative lambdas, one thus obtains an inverse pattern for volatility- and risk premium impulse response functions. To conclude, one can state that there is empirical evidence for time variation in lambdas. Important questions remain such as the partially negative estimated lambdas that are not in line with the CAPM. We consider this as a new area of research, and more empirical work dealing with other stock markets and longer time periods needs to be done.

References

- Baba, Y., R.F. Engle, D.F. Kraft, K.F. Kroner (1990), Multivariate simultaneous generalized ARCH, *mimeo*, Department of Economics, University of California, San Diego.
- Baillie, R.T., T. Bollerslev, H.O. Mikkelsen, (1996), Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 74, 3–30.

- Black, F. (1976), Studies in stock price volatility changes, *Proceedings of the 1976 Meeting of the Business and Economics Statistics Section*, American Statistical Association, 177–181.
- Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307–327.
- Bollerslev, T., R.F. Engle, J.M. Wooldridge (1988), A capital asset pricing model with time varying covariances, *Journal of Political Economy*, 96, 116–131.
- Bollerslev, T., J.M. Wooldridge (1992), Quasi–maximum likelihood estimation and inference in dynamic models with time–varying covariances, *Econometric Reviews*, 11, 143–172.
- Braun, P.A., D.B. Nelson, A.M. Sunier (1995), Good News, Bad News, Volatility, and Betas, *Journal of Finance*, 50, 1575–1603.
- Engle, R. (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica*, 50, 987–1008.
- Engle, R.F., K.F. Kroner (1995), Multivariate simultaneous generalized ARCH, *Econometric Theory*, 11, 122–150.
- Engle, R.F., D.M. Lilien, R.P. Robins (1987), Estimating time varying risk premia in the term structure: The ARCH-M model, *Econometrica*, 55, 391–407.
- Engle, R.F., V.K. Ng (1993), Measuring and testing the impact of news on volatility, *Journal of Finance*, 48, 1749–1778.
- Gallant, A.R., P.E. Rossi, G. Tauchen (1993), Nonlinear dynamic structures, *Econometrica*, 61, 871–907.
- Glosten, L.R., R. Jagannathan, D.E. Runkle (1993), On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, 48, 1779–1801.
- Gouriéroux, C. (1997), *ARCH Models and Financial Applications*, Springer Series in Statistics, Springer Verlag, New York.
- Hafner, C.M., H. Herwartz (1998a), Structural analysis of portfolio risk using beta impulse response functions, *Statistica Neerlandica*, 52, 336–355.
- Hafner, C.M., H. Herwartz (1998b), Volatility impulse response functions for multivariate GARCH models, CORE Discussion Paper 9847, Louvain-la-Neuve.

- Hafner, C.M., H. Herwartz (1999), Testing for linear autoregressive dynamics under heteroskedasticity, SFB 373 Discussion Paper, Humboldt–Universität zu Berlin.
- Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl, T.C. Lee (1988), *Introduction to the Theory and Practice of Econometrics*, Wiley, New York.
- Koop, G., M.H. Pesaran, S.M. Potter (1996), Impulse response analysis in nonlinear multivariate models”, *Journal of Econometrics*, 74, 119–147.
- Lintner, J. (1965), The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics*, 47, 13–37.
- Lumsdaine, R.L. (1996), Consistency and asymptotic normality of the quasi maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, *Econometrica*, 64, 575–596.
- Sharpe, W.F. (1964), Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance*, 19, 425–424.
- Zakoian, J.M. (1994), Threshold Heteroskedastic Functions, *Journal of Economic Dynamics and Control* 18, 931–955.

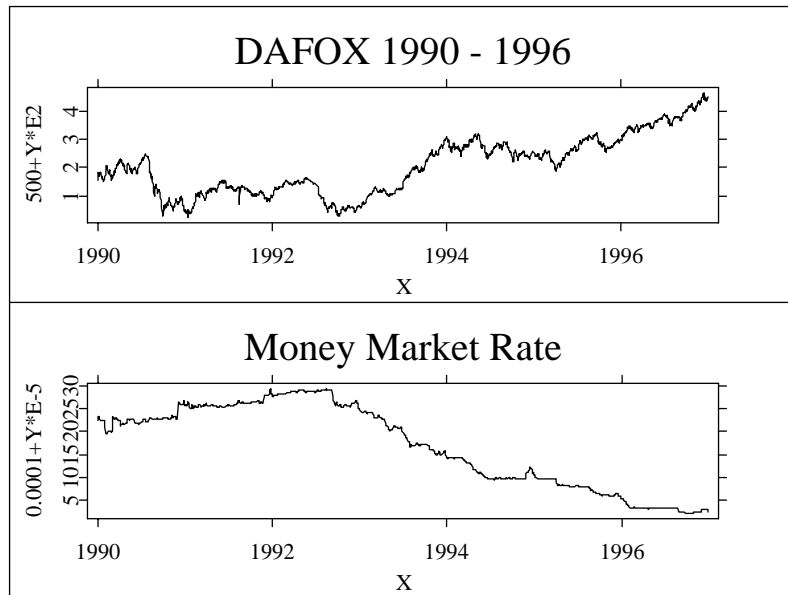


Figure 1: *The upper plot shows the DAFOX series for the period 1990 to 1996. The lower plot shows the German money market rate for the same period.*

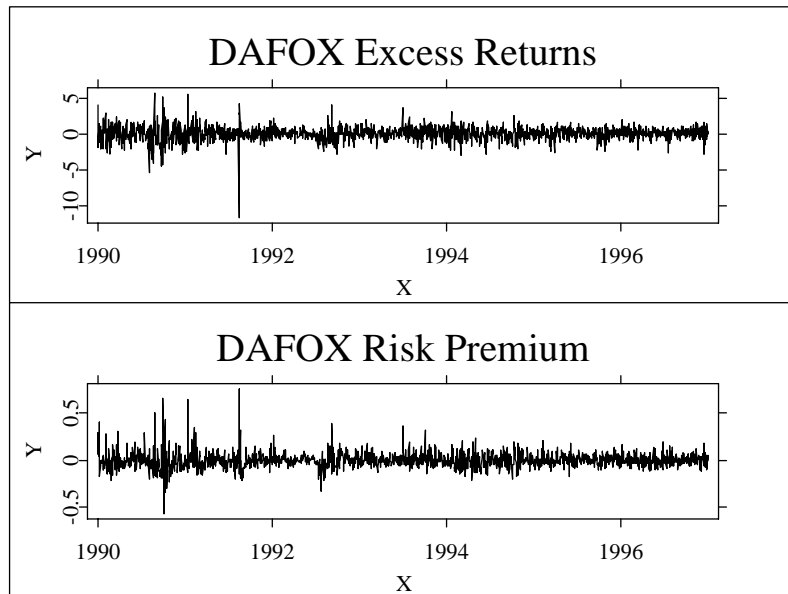


Figure 2: *The upper plot shows the DAFOX excess returns for the period 1990 to 1996. The lower plot shows the estimated conditional mean, i.e. the risk premium.*

	$B = 0$	$2\Delta \log L$	M1	M2	M3	M4
ALLIANZ	3579.9	29.4	3628.8	3578.8	3588.0	3565.3*
BASF	3871.1	38.8	3856.7	3851.8*	3852.9	3851.8
BAYER	3790.5	49.8	3782.2	3765.6*	3776.3	3762.6
BAY.VB	3996.5	48.9	4030.8	4008.2	3981.3	3972.0*
COM. BK.	3841.9	32.2	3873.8	3826.1*	3862.3	3825.8
DAIMLER	3680.1	27.1	3673.1	3666.9	3666.8*	3666.7
DT. BK.	3484.2	40.8	3501.8	3475.3	3476.3	3463.8*
DRSD.BK.	3814.5	51.9	3811.5	3790.1	3788.2*	3788.5
HENKEL	4370.7	10.6	4413.7	4392.3	4366.2*	4365.4
HOECHST	3991.1	32.2	3985.1	3975.9*	3979.5	3975.0
KARSTADT	4293.2	27.8	4311.0	4300.1	4281.8*	4279.3
LINDE	4212.8	21.2	4223.9	4208.4	4208.7	4202.2*
MAN	4125.2	35.1	4157.4	4106.9	4118.9	4101.7
MANNESM	4047.9	13.4	4093.5	4041.9	4041.2*	4037.5
M. RUECK.	4400.9	24.4	4421.4	4394.3	4396.0	4388.7*
PREUSSAG	4115.2	23.6	4121.2	4103.6*	4104.7	4105.3
RWE	3777.9	40.4	3831.9	3770.8	3799.7	3757.7*
SIEMENS	3510.6	57.8	3494.1	3482.5	3481.7*	3480.4
THYSSEN	4069.1	31.0	4081.0	4053.6*	4057.4	4052.8
VIAG	4180.3	11.8	4192.8	4174.7	4174.4*	4171.4
VW	3882.2	27.6	3879.0	3873.2	3868.4*	3867.8

Table 1: *Estimation results (negative log-likelihood) for alternative volatility models under the assumption of time invariance of λ_t . The selected model is indicated by an asterisk. Restricted versions ($B = 0$) of the selected model were also estimated. $2\Delta \log L$ denotes two times the log likelihood difference between the unrestricted and restricted model.*

	$-\log L$	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
ALLIANZ	3565.3	1.15	8.12*	6.35*	2.68	0.07	8.04*	3.09	2.25	0.03	8.73*	4.71	2.70
BASF	3851.8	0.54	7.10*	1.99	2.54	3.26	10.84*	3.29	3.48	1.00	12.33*	3.17	2.94
BAYER	3765.6	1.07	2.11	1.11	0.41	-0.20	4.58*	3.44	2.27	-0.84	3.88	6.05*	4.08
BAY. VB.	3972.0	4.13*	1.49	2.04	1.69	3.65	6.67*	6.58*	1.80	3.60	5.02	11.04*	2.50
COM. BK.	3826.1	6.05*	4.23*	4.83*	5.87*	6.09*	6.53*	3.75	6.86*	4.88	4.46	6.30*	5.82
DAIMLER	3666.8	2.12	0.43	0.73	0.83	3.87	4.67*	5.04*	2.83	4.17	9.67*	12.85*	8.31*
DT. BK.	3463.8	-0.15	1.96	1.33	1.77	-0.62	0.06	-0.41	-0.59	-1.27	3.32	2.35	2.27
DRSD. BK.	3788.2	2.71	2.40	1.84	2.62	0.45	5.49*	3.89	3.58	0.40	7.62*	7.72*	4.65
HENKEL	4366.2	-0.99	1.21	0.33	0.31	-1.51	2.83	0.96	1.02	-1.06	2.43	0.73	0.82
HOECHST	3975.9	1.51	0.56	0.88	1.30	0.24	1.41	-0.41	0.17	-0.80	3.53	2.34	1.36
KARSTADT	4281.8	6.72*	1.97	1.75	3.09	3.66	3.66	2.66	3.85	3.66	1.53	3.01	3.30
LINDE	4202.2	1.65	1.50	0.99	0.83	0.62	2.90	2.06	2.61	0.88	5.27	1.66	4.69
MAN	4101.7	-0.39	0.14	0.17	0.60	1.22	0.48	1.01	0.84	-0.27	1.66	2.65	0.63
MANNESM	4041.2	0.09	1.98	0.91	0.72	-0.03	3.94	2.42	2.42	-0.33	4.56	3.51	2.74
M. RUECK.	4388.6	0.00	3.71	1.80	1.44	0.71	5.85*	1.79	1.34	0.13	4.66	1.84	1.44
PREUSSAG	4103.6	1.98	5.80*	2.91	3.54	2.24	6.35*	3.20	3.45	1.98	6.83	3.13	3.65
RWE	3757.7	5.50*	0.01	0.77	2.62	3.92	3.07	2.89	4.06*	1.19	0.12	0.87	3.13
SIEMENS	3481.7	2.02	1.87	1.59	2.24	0.14	3.77	3.23	2.62	0.17	4.77	2.37	2.74
THYSSEN	4053.6	0.86	2.28	2.51	0.95	-0.75	3.66	4.41*	2.42	-0.10	4.47	3.67	0.80
VIAG	4174.4	5.75*	2.74	1.89	3.44	1.63	5.46*	3.10	4.95*	1.46	4.51	3.20	4.24
VW	3868.4	0.23	1.11	0.81	0.91	0.10	1.72	2.04	1.02	0.15	4.05	3.68	1.03

Table 2: Estimation results for alternative specifications of λ_t . Entries for L1 to L12 indicate twice the difference of the log-likelihood function for the respective model with time dependent λ_t relative to the 'best' specification assuming $\lambda_t = \lambda$.

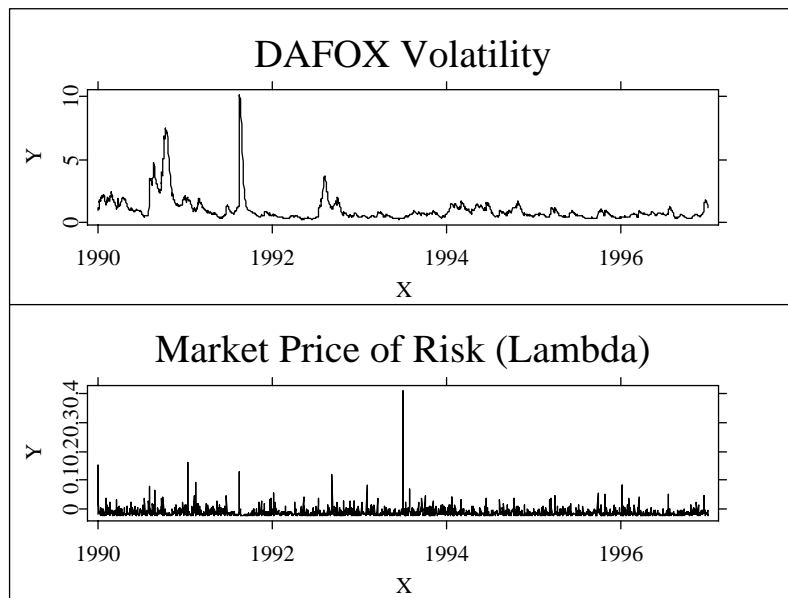


Figure 3: *The upper plot shows the estimated volatility for the DAFOX. The lower plot shows the estimated market price of risk, λ_t . Both series were obtained by estimation of (17) for the bivariate series DAFOX-ALLIANZ.*

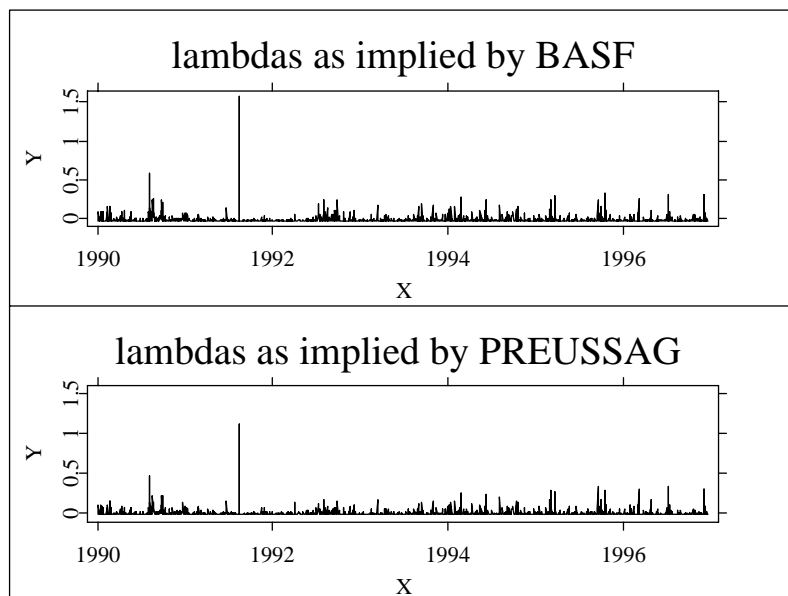


Figure 4: *The estimated market price of risk (lambda) as implied by the estimation of (17) for the bivariate series DAFOX-BASF (upper plot) and DAFOX-PREUSSAG (lower plot).*

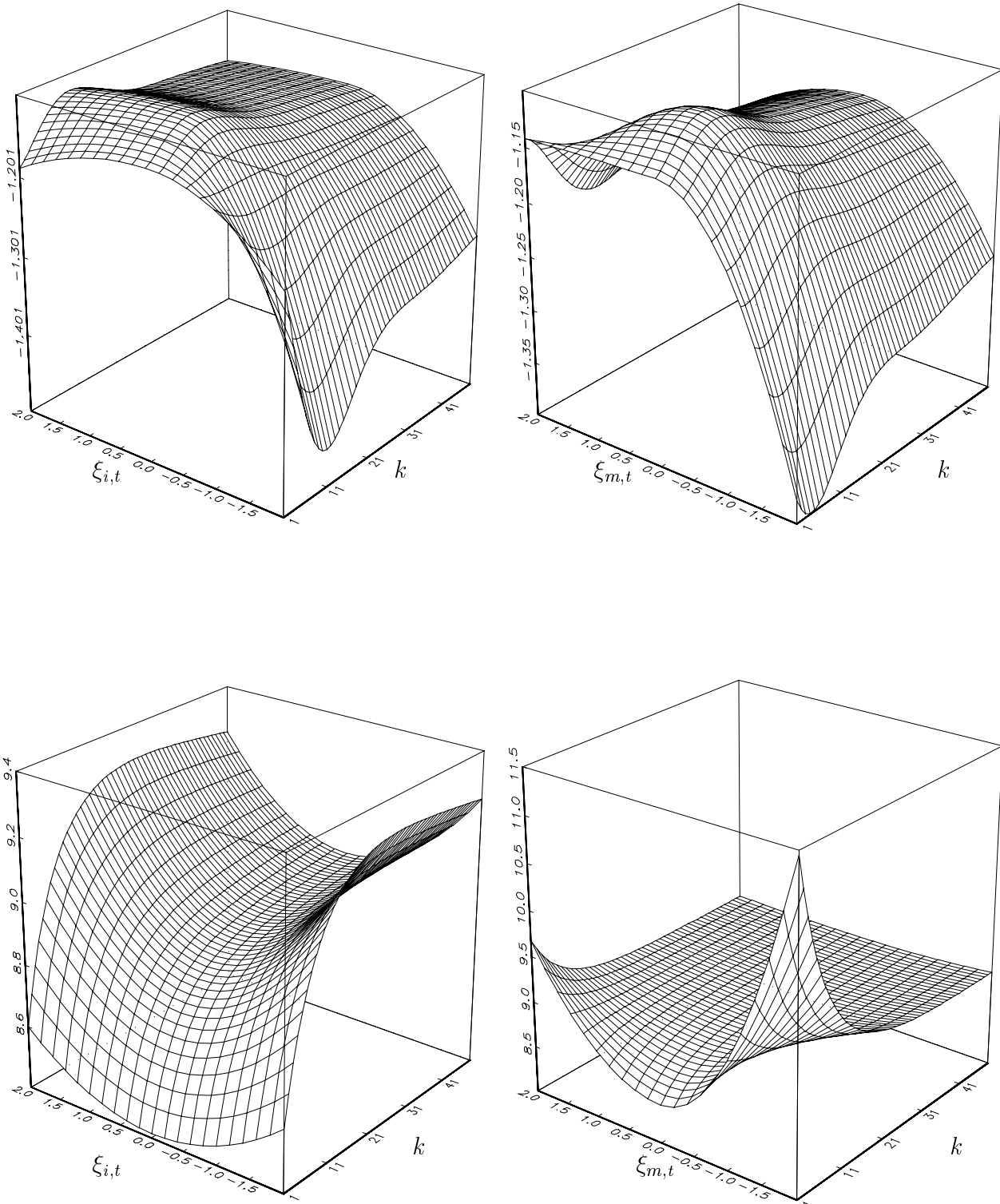


Figure 5: Estimates of the impulse response functions $P_k(\xi_t)$ as defined in (19) for the risk premium $(\lambda_t \sigma_{m,t}^2)$ based on the bivariate GARCH-M model ALLIANZ-DAFOX (upper panel, model L2) and DAIMLER-DAFOX (lower panel, model L6). The independent innovations are asset specific (left panels) and innovations to the DAFOX. The right axes indicate the evolution over time up to 50 periods. The scale of all ordinates is E-02.