About the Impact of Model Risk on Capital Reserves: A Quantitative Analysis.

Philip Bertram[‡1], Philipp Sibbertsen[‡] and Gerhard Stahl[§]

‡Leibniz University of Hanover, Department of Economics, Germany
§Talanx AG, Hanover and Ulm University

January 2015

ISSN

Abstract

This paper analyzes and quantifies the idea of model risk in the environment of internal model building. We define various types of model risk including estimation risk, model risk in distribution and model risk in functional form. By the quantification of these concepts we analyze the impact of the modeling process of an econometric model on the resulting company model. Utilizing real insurance data we specify, estimate and simulate various linear and nonlinear time series models for the inflation rate and examine its impact on pension liabilities under the aspect of model risk. Under consideration of different risk measures it is shown that model risk can differ profoundly due to the specification process of the econometric model resulting in substantial variations of capital reserves. We further compare our definition of model risk to the standard Basel approach interpreting model risk as a constant multiplication factor with regard to market risk. We show that these different definitions of model risk can lead to remarkable monetary differences concerning induced capital reserves.

JEL codes: G12, G18

Keywords: Model risk, Estimation risk, Misspecification risk, Basel multiplication factor, Empirical model specification, Capital reserves.

[1]corresponding author:
phone: +49- 511- 762-3082, fax: +49- 511- 762-3923
e-mail: bertram@statistik.uni-hannover.de

We gratefully acknowledge the financial support of the Fritz-Thyssen-Stiftung. Further we thank two anonymous referees as well as the editor for helpful comments and suggestions.

1 Introduction

From a financial institution's point of view the importance of dealing with model risk has risen substantially since the implementation of new regulatory laws such as Basel II or Solvency II. Since then the option of implementing internal models instead of the hitherto obligatory application of standard methods as e.g. being documented in QIS 4b for the calculation of the solvency capital requirement has been driven forth. Internal models are particulary suitable for covering the risen demands of stakeholders concerning the quality of risk management as the incorporation of sophisticated and flexible mathematical methods can be fulfilled. Another advantage of internal models apart from the improved risk measurement marks the refinement of the risk culture. This might be exemplified by the procedure of rating agencies demanding the existence of an internal model in order for the company to be rated *strong* concerning its risk management. Internal models can be defined as large (high amount of explanatory variables), nonlinear (embedded options), stochastic (modeling future states of nature) systems. In the context of the holistic approach of Basel II and Solvency II an estimation of the balance sheet's forecast distribution is carried out by consulting company models (management rules, provision for premium refunds etc.) as well as stochastic models.

Nevertheless the implementation of internal models implies one thus far not satisfactorily handled issue: the topic of model risk. Without consideration of the latter the capital reserves are determined by the standard approach of risk management. That is portfolio risk is subsumed as the aggregate of the marginal distributions of the risk factors market risk, credit risk and operational risk applying a suitable aggregation method (for a discussion of this topic cf. Rosenberg and Schuermann [2006]) and reporting a risk measure thereof. With the possible utilization of internal models in order to model market risk the risk measure of the latter depends substantially on the concrete specification of the internal model. Thus there does exist a strong relationship between model risk and the resulting market risk which should be accounted for when it comes to the determination of capital reserves. In this context we understand model risk as every risk induced by the choice, specification and estimation of a statistical model.[2]

In order for model risk to be considered as a separate risk factor an operational quantification of the former should be provided. Although some authors like Crouhy et al. [1998] or Cont [2006] made several proposals for an abstract coverage of the topic there does not exist an unambiguous method for the quantification of model risk thus far. In the literature there are basically two approaches dealing with the question of measuring model risk: the bayesian model averaging approach (cf. e.g. Brock et al. [2003]) and the worst-case approach (cf. e.g. Kerkhof et al. [2010]). Although from a practical point of view there is no such thing as obligatory capital charges for induced model risk the Basel Committee (cf. BCBS [1996]) suggests a so-called multiplication factor of three with regard to market risk in order to account for model risk. Stahl [1997] showed that the multiplication factor may be interpreted as the relation of the risk measure of the underlying under different (parametric or non-parametric) distributions. This interpretation corresponds closely to the worst-case approach of measuring model risk. Hence in this paper we follow the idea of Kerkhof et al. [2010] fragmenting model risk into estimation risk, misspecification risk and identification risk and analyze its impact on capital reserves. Our approach features the following new aspects concerning the topic of model risk and capital reserves.

[2] Note that human failure is captured under operational risk.

By using real insurance data we do not only analyze the model risk of the underlying but also take the company model into account what, to our knowledge, has not been done before. Whilst the existing literature does not differentiate between the statistical model and the company model resulting in the assumption that the underlying marks a concrete balance sheet position we take the whole structure of the internal model into account. Concretely we utilize a specific company model, the model for pension liabilities of a large German insurance company, and demonstrate its interaction with the econometric model for the underlying under the aspect of model risk. By taking a broad range of time series models into account which differ in their functional forms we are able to refine the definition of model risk further by discriminating between misspecification risk in functional form and misspecification risk in distribution and are thus enabled to quantify its contribution to overall model risk.

The paper proceeds as follows. A formal definition of the various types of model risk is carried out in section 2. Afterwards the results of the empirical study are presented in section 3. Here we briefly describe the specification and estimation results for the underlying first. Then we look at the implications concerning model risk with respect to the underlying and the company model. Section 4 concludes.

2 Measuring Model Risk

2.1 Econometric Setting

One of the main tasks of risk management is to determine a risk measure, denoted by π, of an economic or financial variable of interest, X which is distributed according to some density function f, short $X \sim f$. Obviously π may then be written as some function f, i.e. $\pi(f)$. Throughout the paper we specify π as the value-at-risk (VaR) being defined as $\pi(p) = \inf\{x \in \mathbb{R} | \mathbb{P}(X \leq x) \geq p\}$[3] where p denotes the confidence level. The *VaR* may then e.g. be interpreted as the $p-$quantile of a loss distribution of a financial asset.

In case f is known π marks the market risk of the underlying X. In practice however f is unknown in two respects. First only a finite number of values for the random variable X can be observed. In other words the data $\{x_t\}_0^T$ with time index $t = (0, \ldots, T < \infty)$ are assumed to be realizations originating from the stochastic process $\{X_t\}_{-\infty}^{\infty}$, i.e. $\{x_t\}_0^T$ is an approximation for the unknown process $\{X_t\}_{-\infty}^{\infty}$. The second part that has to be approximated in $\pi(f)$ concerns the character or shape of the density function f which is also unknown in practice. Denoting the density function under an assumed candidate model i by f_i leads to the potential risk that $f \neq f_i$ and thus $\pi(f) \neq \pi(f_i)$. These two concepts define the two components of total model risk: estimation risk and misspecification risk.

In order to overcome the gap between $\{x_t\}$ and $\{X_t\}$ the frequentist's approach assumes the data to be generated by a so-called data generating process (DGP). The DGP connects the theoretical distribution of X with its empirical counterpart by introducing the k-dimensional parameter space $\Theta \subseteq$

[3] We additionally utilized the expected shortfall being proposed by Artzner et al. [1999]: $ES(p) = E_{\mathbb{P}}(x \in \mathbb{R} | x \geq \inf\{x \in \mathbb{R} | \mathbb{P}(X \leq x) \geq p\})$ as a risk measure. Since the results do not differ qualitatively we solely report the results for the value-at-risk. Quantitative results for the expected shortfall are available upon request.

$\mathbb{R}^k$ which determines the character of the empirical distribution function. The population parameter $\theta \in \Theta$ marks the point in Θ that generated the data. The generic econometric DGP ist given by the model

$$x_t = H(z_t|\theta) + \varepsilon_t \quad \text{with} \quad \varepsilon_t \sim G \quad \text{and} \quad \varepsilon_t \perp z_t. \tag{2.1}$$

$z_t = (x_{t-1}, \ldots, x_0, y_t, \ldots, y_0) = (\tilde{x}_{t-1}, \tilde{y}_t)$ contains all kinds of lagged endogenous $(\tilde{x}_{t-1})$ and/or exogenous $(\tilde{y}_t)$ explanatory variables, $H(\cdot)$ describes the functional form of the relationship and (ε_t) denotes an error term process being distributed according to distribution function G. We further impose (ε_t) to be identically but not necessarily to be independently distributed. With the assumption of no endogeneity in 2.1 leads to f being given as

$$f(x) = \int_{-\infty}^{\infty} H(x-\varepsilon)G(\varepsilon)d\varepsilon. \tag{2.2}$$

2.2 Definitions of Model Risk

Let us now introduce the empirical counterpart of 2.2 as being given as

$$f_i(x) = f_{jk}(x) = \int_{-\infty}^{\infty} H_j(x-\varepsilon)G_k(\varepsilon)d\varepsilon. \tag{2.3}$$

The empirical or assumed model is defined as M_i implying that the data is distributed according to f_i. Then the mapping $j \oplus k \to M_i$ returns the empirical or assumed model specification. In other words the specification of a functional form j and an error distribution assumption k where $j \in \mathcal{J}$ and $k \in \mathcal{K}$ denote sets of feasible functional form and distribution specifications are combined such that through a combination of j and k a model set $i \in I$ is determined. In other words every combination of a specification of functional form j and distribution k leads to the concrete model specification i. 2.2 and 2.3 are assumed to be related via $f_{ab} = f$. Hence choosing $k = a$ and $j = b$ yields the true density specification of x which practically is obviously never possible as a and b cannot be observed. This enables us to analyze the topic of model risk according to specification i with respect to four different assumptions concerning the choice of j and k in 2.3:

$$\begin{aligned}
A_1 &: \quad a = j \quad \text{and} \quad b = k \\
A_2 &: \quad a \neq j \quad \text{and} \quad b \neq k \\
A_3 &: \quad a \neq j \quad \text{and} \quad b = k \\
A_4 &: \quad a = j \quad \text{and} \quad b \neq k
\end{aligned}$$

Example 1. *Let the random variable X follow an autoregressive model of order one ($AR(1)$ model) being defined as $X_t = \beta \cdot X_{t-1} + \varepsilon_t$ with time index t and error term process $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$. Let further* *be the* $AR(1)$ *parameter* β *have the two possible specifications* $(0.1, 0.5)$ *and the variance of* ε *be either 1 or 0.5. Then*

A_1 : β *and* σ^2 *are specified correctly*

A_2 : *Neither* β *nor* σ^2 *are specified correctly*

A_3 : β *is not specified correctly and* σ^2 *is specified correctly*

A_4 : β *is specified correctly and* σ^2 *is not specified correctly.*

These four assumptions are now referred to for defining model risk in an operational way. Let the three resulting density functions for x according to model specification i with respect to assumptions $A1-A4$ be defined as $f_{i,A_1} = f_{jk}(x, a=j, b=k)$, $f_{i,A_2} = f_{jk}(x, a \neq j, b \neq k)$, $f_{i,A_3} = f_{jk}(x, a \neq j, b = k)$, $f_{i,A_4} = f_{jk}(x, a=j, b \neq k)$ whereas $f_{jk} = f_i$ describes the unconditional empirical density of X.

The connection of the density functions under the four assumptions and the resulting risk measure concerning model specification i is described by $\pi_i = q(f_i)$. Since π_i marks a point estimate of a quantile of f_i a confidence interval $[\pi_i \pm \eta_i(\alpha)]$ for the estimate at level $1-\alpha$ can be defined.[4] The upper bound of the confidence interval marks a crucial part in our model risk definitions. As estimation risk appears under every possible model we account for this by considering the risk of an estimation error in the VaR. Via the functional delta method η can be derived as $\eta_{iT}(\alpha) = z_{1-\alpha/2}\hat{\sigma}_i$[5], where

$$\hat{\sigma}_i = \sqrt{p(1-p)/(T \cdot f_i(\pi_i(p))^2)} \tag{2.4}$$

with p denoting the risk measure's confidence level and $z_{1-\alpha/2}$ being the $(1-\alpha/2)-$quantile of the standard normal distribution. Hence $\hat{\sigma}_i$ marks an estimate of the standard deviation of the estimate π_i. As the sample size increases the standard deviation goes to zero, i.e. $\lim_{T\to\infty} \eta = 0$ as via (2.4) $\lim_{T\to\infty} \hat{\sigma}_i = 0$. As the standard deviation is asymptotically normal, (2.4) is multiplied by the respective quantile of the standard normal distribution $z_{1-\alpha/2}$. The corresponding conditional values are denoted by π_{i,A_γ} and η_{i,A_γ} with $\gamma \in \{1,2,3,4\}$ where i, A_γ indicates that it is referred to f_{i,A_γ} in the respective formula. π and σ describe the point estimate and standard deviation with respect to the true density function $f(\cdot)$.

Under A_1 $f_i = f_{i,A1} = f$ is not misspecified, i.e. f is modeled correctly which is why there is no misspecification risk under A_1. There is, however, estimation risk as $f_{i,A_1}(x) \neq f$ as $T < \infty$ in $\{x_t\}_0^T$ with T denoting the sample size. Consequently the quotient of the upper bound of the confidence interval and the point estimate under A_1 yields an operational definition for estimation risk according to model specification i.

Definition 1 (Estimation Risk). *Let* $\Pi_{i,A_1} = \pi_{i,A_1} + \eta_{i,A_1}$ *be the upper bound of the* $1-\alpha$ *confidence interval of the risk measure's point estimate according to model specification i under the assumption*

[4] Throughout the paper we specify α to the standard value of 0.05. We also chose values of 0.1 and 0.01 yielding qualitatively the same results.

[5] For notational convenience we drop the index T in the following.

that $a = j$ and $b = k$ in (2.3). Estimation risk of model i is then given by $R_{1i}(\alpha; p; T) = \Pi_{i,A_1} \cdot \pi_{i,A_1}^{-1} = 1 + \eta_{i,A_1} \cdot \pi_{i,A_1}^{-1}$.

Example 2. *Let X be given as described in example 1. Let us further denote the specification of* $(\beta, \sigma^2) = (0.5, 1)$ *as model i. Then* f_{i,A_1} *is obviously given as* $f_{i,A_1} = N(0, (1-0.5^2)^{-1})$*. With* $\alpha = 0.05$, $p = 0.995$ *and* $T = 1000$*, estimation risk of model i under the assumption that model i is correctly specified is given as*

$$
\begin{aligned}
R_{1i}(0.05; 0.995; 1000) &= 1 + \frac{\eta_{i,A_1}}{\pi_{i,A_1}} = 1 + \frac{1.96 \cdot \sqrt{0.995 \cdot 0.005 / (1000 \cdot f_{i,A1}(2.97)^2)}}{2.97} = \frac{3.3234}{2.97} \\
&= 1.1190.
\end{aligned}
$$

In this definition and throughout the paper model risk is measured in ratios. This facilitates a comparison between the risk definitions of different economic variables measured in different units. Model risk is then interpreted as the the error that is made under the (wrongly) imposed model specification compared to the point estimate of the risk measure if this error had not been made. If e.g. a specific functional form is imposed and model risk in distribution is smaller than 1 it can be interpreted that the upper bound of the confidence interval of the model under the imposed distribution is still smaller than the point estimate under the correct distribution. Hence this may occur when e.g. a t-distribution with 20 degrees of freedom is assumed while 5 is the correct number.

Further it should be mentioned that in the above definition of model risk the latter is not understood as model risk inherent in a product but rather as a relative measure. Precisely, total model risk can be interpreted as excess market risk. While market risk is covered by the fluctuation of the underlying factor and hence by the determination of π, the "excess" is captured by the risk which arises through estimation and misspecification risk. Altogether total model risk may then function as a capital charge that covers risk sources beyond the common market risk definition.

The second component of model risk arises from the fact that f is unknown in practice. Hence practically A_1 can never be shown to be fulfilled. Via the data process $\{x_t\}_0^T$ where $T < \infty$, f can be approximated via some nonparametric kernel density estimate of X under model i. We call this density estimate the empirical density function and denote it by $\tilde{f}$.[6] The distance from the empirical density to the true density is then given by $|\tilde{f}(x) - f|$.

The consistency is shown by the Glivenko–Cantelli[7] theorem for *iid* processes as

$$
P\left(\lim_{T\to\infty} \sup_x |\tilde{f}(x) - f| = 0\right) = 1.
$$

Therefore the empirical density function under model i is seen as an approximation for f resulting in the risk of misspecifying the population density function when j and/or k are chosen incorrectly. On top there is again an estimation uncertainty even when f is found as described earlier. Therefore, total model risk concerning model i consists of estimation risk and misspecification risk and can be defined as the difference of π regarding the true model and the imposed model specification i.

[6]Although the empirical density estimate depends on the sample size T we drop the index for simplicity.

[7]Tucker [1959] derives a generalization of the theorem for stationary time series and thus also covers all DGPs in (2.1).

The definition of total model risk can then be formulated as follows.

Definition 2 (Total Model Risk)**.** *Let* $\Pi_{A_1} = \pi_{A_1} + \eta_{A_1}$ *be the upper bound of the* $1-\alpha$ *confidence interval of the risk measure's point estimate under the assumption that* $a = j$ *and* $b = k$ *in (2.3) and let* π_{i,A_2} *be the point estimate of* π *under the assumption that* $a \neq j$ *and* $b \neq k$ *in (2.3). Total model risk of model* i *is then given by* $R_{2i}(\alpha; p; T) = \Pi_{A_1} \cdot \pi_{i,A_2}^{-1}$.

Example 3. *Let us now falsely assume that* $(\beta, \sigma^2) = (0.1, 0.5)$ *and let this specification define model* i *whereas* $(\beta, \sigma^2) = (0.5, 1)$ *are the true values. As then obviously* $\pi_{i,A_2} = 1.83$ *and with* $\alpha = 0.05$, $p = 0.995$ *and* $T = 1000$ *total model risk is given as*

$$R_{3i}(0.05; 0.995; 1000) = \frac{\pi_{i,A_1} + \eta_{i,A_1}}{\pi_{i,A_2}} = \frac{3.3234}{1.83} = 1.8161.$$

The inequality $f_i \neq f$ holds when either the functional form assumption is wrong, i.e. $j \neq a$ and/or the error distribution is modeled incorrectly, i.e. $k \neq b$. Consequently misspecification risk may be further differentiated into misspecification risk in functional form and misspecification risk in distribution depending on whether A_3 or A_4 is met. Note however that as $T < \infty$ in empirical modeling there is still estimation risk. Thus the term misspecification risk may be misleading as rather a combination of estimation risk and misspecification risk is defined. Straightforwardly the two types of misspecification risk may be formulated.

Definition 3 (Misspecification Risk in Functional Form)**.** *Let* $\Pi_{A_1} = \pi_{A_1} + \eta_{A_1}$ *be the upper bound of the* $1-\alpha$ *confidence interval of the risk measure's point estimate under the assumption that* $a = j$ *and* $b = k$ *in (2.3) and let* π_{i,A_3} *be the point estimate of* π *under the assumption that* $a \neq j$ *and* $b = k$ *in (2.3). Total model risk of model* i *is then given by* $R_{3i}(\alpha; p; T) = \Pi_{A_1} \cdot \pi_{i,A_3}^{-1}$.

Thus R_{3i} returns a measure of what happens when A_3 is imposed whilst A_1 is true. In other words the effect of falsely imposing $j = a$ whereas in truth $j \neq a$ holds true is quantified under the condition that the error distribution is specified correctly.

Remark 1. R_{3i} *may likewise be defined as the risk of the functional form being misspecified under a wrongly imposed distribution specification, i.e.* $R_{3i}(\alpha; p; T) = \Pi_{A_4} \cdot \pi_{i,A_2}^{-1}$.

Example 4. *Example 3 is now modified such that model* i *is specified such that* $(\beta, \sigma^2) = (0.1, 1)$ *resulting in* $\pi_{i,A_3} = 2.59$*. With* $\alpha = 0.05$, $p = 0.995$ *and* $T = 1000$ *total model risk is given as*

$$R_{3i}(0.05; 0.995; 1000) = \frac{3.3234}{2.59} = 1.2832.$$

Under the equivalent form

$$R_{3i}(0.05;0.995;1000) = \Pi_{A_4} \cdot \pi_{i,A_2}^{-1} = \frac{2.10 + 1.96 \cdot \sqrt{0.995 \cdot 0.005/(1000 \cdot f_{i,A4}(2.10)^2)}}{1.83} = 1.2832.$$

Straightforwardly comparing π under A_1 and A_4 yields a definition for misspecification risk in distribution.

Definition 4 (Misspecification Risk in Distribution)**.** *Let $\Pi_{A_1} = \pi_{A_1} + \eta_{A_1}$ be the upper bound of the $1-\alpha$ confidence interval of the risk measure's point estimate under the condition that $a = j$ and $b = k$ in 2.3. Misspecification risk in functional form of model i is then given by $R_{4i}(\alpha; p; T) = \Pi_{A_1} \cdot \pi_{i,A_4}^{-1}$.*

Remark 2. *R_{4i} may likewise be defined as the risk of the distributional form being misspecified under a wrongly imposed functional form specification, i.e. $R_{4i}(\alpha; p; T) = \Pi_{A_3} \cdot \pi_{i,A_2}^{-1}$.*

Example 5. *Let finally model i be specified such that $(\beta, \sigma^2) = (0.5, 0.1)$ resulting in $\pi_{i,A_4} = 2.10$. With $\alpha = 0.05$, $p = 0.995$ and $T = 1000$ total model risk is given as*

$$R_{4i}(0.05;0.995;1000) = \frac{3.3234}{2.10} = 1.5826.$$

Under the equivalent form

$$R_{4i}(0.05;0.995;1000) = \Pi_{A_3} \cdot \pi_{i,A_2}^{-1} = \frac{2.59 + 1.96 \cdot \sqrt{0.995 \cdot 0.005/(1000 \cdot f_{i,A3}(2.59)^2)}}{1.83} = 1.5826.$$

So far solely the underlying economic variable has been considered. In practice, however, the variables are often modeled as an input variable for a so-called company model. This company model can e.g. be a pension model of an insurance company (cf. section 3) and imposes itself a functional form according to which the economic variables enter the model. Hence the variable at interest (e.g. pensions denoted by L) depend on the underlying via the relationship $P = g(X)$ where $g(\cdot)$ denotes a continuously differentiable function.[8] Then the density of P with respect to f_i is given by

$$\hat{f}_i(P) = f_i \cdot |dh(P)/dP| \text{ where } h(P) = g^{-1}(P). \tag{2.5}$$

Hence the various definitions of market risk and model risk (total, in distribution and in functional form) can easily be transferred to the company model by substituting by applying equation (2.5). Note, however, that we cannot define an estimation error in this setting since the company model cannot be handled like an econometric model.

[8] A concrete specification of $g(\cdot)$ is given in section 3.

3 Empirical Study

In this section we show that the definition of model risk is not only an academic exercise but that it marks a highly relevant topic in practice due to its monetary implications. Concretely we exemplify the monetary changes concerning the capital reserves from an internal company model that occur when using different definitions of model risk. For this purpose we use a real company model stemming from a large German insurance company. Concretely we deal with the company's pension model and its implied capital reserves. The pension liabilities depend on two economic variables: inflation and interest rates. Whereas the former functions as an adjustment for the obligation in terms of salary rates the latter is used as a discount rate. The pension liabilities are then given by

$$P = \beta_0 \cdot (1 + \beta_1 \cdot I)^{\beta_2} \cdot (1 + \beta_3 \cdot Y)^{-\beta_2}, \tag{3.1}$$

where I denotes inflation, Y denotes the interest rate and $\beta = (\beta_0, \ldots, \beta_3)$ returns a vector of coefficients. Due to reasons of concealment β cannot be reported in this paper. Fig.3 however gives an idea of the shape of the function. In order to derive the distribution of the pension liabilities from

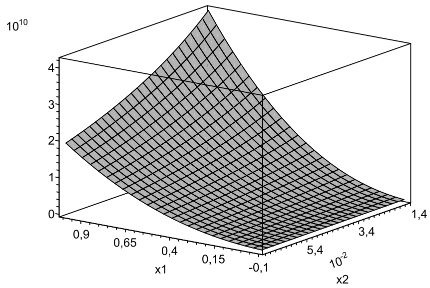

Figure 1: illustrates the shape of the pension function. x_1 describes the inflation rate, the interest rate is given by x_2 and P is displayed on the z-axis. Note that the combination of high inflation rates and low interest rates leads the function to rise quickly.

which the capital reserves are determined different inflation and interest rate scenarios have to be considered. Concretely we simulated $J = 10,000$ inflation scenarios for $I = 15$ different specifications of various econometric time series models. Following we calculated $P_{j,i}$ according to 3.1 with $j = (1, \ldots, J = 10,000)$ and $(i = 1, \ldots, I = 15)$ resulting in the array of pension liabilities $P \sim (J \times K)$. The concrete specification of the inflation models is described in the next section.

3.1 Inflation Models

In order not to further complicate the procedure we solely deal with modeling the inflation rate for a start. As far as the interest rate is concerned scenarios having been developed internally by the insurance company are utilized (for a brief overview of the interest rate scenario's distribution cf. panel (a) in Fig.2). This is legitimized by considering the position of the inflation rate at the top of the cascade in the benchmark Wilkie model (Wilkie [1995]) mirroring its particular importance. In the model of Wilkie [1995] ecomic variables are modeled univariately and put together in a cascade model. Concerning Inflation for example firstly an $AR(1)$ model is fitted while secondly scenarios based on the fitted model are simulated. Those scenarios then function as an input for another economic variable such as interest rates.

In case a misspecified inflation model is utilized the misspecification error transmits throughout the whole system. This means that dealing with the inflation model should be of the highest priority when it comes to specifying scenarios that are to outperform the Wilkie model.

In order to carry out a consistent procedure of model specification a hierarchy of the univariate time series models being used in practice is very helpful. In the first level we discriminate between linear and nonlinear models. Recall that by the Wold decomposition any zero-mean purely non deterministic stationary process $\{I_t\}_{t=1}^T$ can be written in the form

$$I_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \tag{3.2}$$

where $\sum_i ||\Psi_i||^2 < \infty$ and $\{\varepsilon_t\}_{t=1}^T$ is a stationary sequence of centered and uncorrelated random variables with common variance Σ. A process $\{I_t\}_{t=1}^T$ is said to be linear when $\{\varepsilon_t\} \overset{iid}{\sim} (0,\Sigma)$ in 3.2. Otherwise the process can be declared nonlinear. Note that the nonlinearity can occur in the mean as well as in the volatility.

This thought leads to the discrimination of three classes of time series models in our model selection procedure: linear models, nonlinear models in the mean and nonlinear models in the volatility.

The linear models are described by the $ARFIMA(p,d,q)$ model class being given as

$$\Theta(L)^{-1}\Phi(L)(1-L)^d I_t = \varepsilon_t, \tag{3.3}$$

where $\{I_t\}_{t=1}^T$ describes the time series of interest and $\{\varepsilon_t\}_{t=1}^T$ forms a white noise process. A possibility to model nonlinearity in the mean in this setting was proposed by Hsu [2005]. By rewriting a time series $I_t = \mu_t + \varepsilon_t$ as the sum of a deterministic part μ_t and a stochastic part ε_t the former can be modelled as $\mu_t = \mu_1 + \sum_{i=1}^n \lambda_i \cdot 1(l_i < t \leq l_{i+1})$ where n denotes the number of breaks, l_i are the break points, $1(\cdot)$ denotes the indicator function and $\lambda_i = \mu_{i+1} - \mu_i$. Note that structural changes in the mean are a typical example of the occurance of spurious long memory (cf. Diebold and Inoue [2001] or Engle and Smith [1999]). By neglecting the mean shifts the estimation of the fractional differencing parameter d might be biased quite heavily. That is why Hsu [2005] proposed to first determine the number of break points in the model and thereafter estimate the $ARFIMA$ parameters and the time of the breaks simultaneously. Whereas the former is done via application of the LIC information criterion described in Lavielle and Moulines [2000] the estimation is carried out by a modified local Whittle method.

A further possibility to model nonlinearity in the mean is given by the $STAR$ model introduced by Chan and Tong [1986] and popularized by Granger and Teräsvirta [1993] and Teräsvirta [1994]. It is given by

$$\begin{aligned} I_t &= (\phi_{0,1} + \phi_{1,1} I_{t-1} + \ldots + \phi_{p_1} I_{t-p_1})(1 - G(I_{t-1};\gamma,c)) \\ &\quad + (\phi_{0,2} + \phi_{1,2} I_{t-1} + \ldots + \phi_{p_2} I_{t-p_2}) G(I_{t-1};\gamma,c) + \varepsilon_t \end{aligned} \tag{3.4}$$

with $G(\cdot)$, γ and c denoting transition function, smoothness parameter and threshold value.

Finally nonlinearities in the volatility may be handled by the *APARCH* model class which was introduced by Ding et al. [1993] and is defined as

$$I_t = \mu + \sum_{i=1}^{p} a_i I_{t-i} + \sum_{j=1}^{q} \varepsilon_{t-j} + \varepsilon_t, \tag{3.5}$$

$$\varepsilon_t = h_t^{1/\delta} \cdot \nu_t,$$

$$h_t = \omega + \sum_{k=1}^{K} \alpha_k \left(|\varepsilon_{t-k}| - \psi_k \cdot \varepsilon_{t-k}\right)^{\delta} + \sum_{l=1}^{L} \beta_l \cdot h_{t-l} \tag{3.6}$$

where μ and ω are constants, a, α and β are vectors of coefficients and $\{\nu_t\} \overset{iid}{\sim} (0,\Sigma)$. The specification and estimation of the models is described in the next section.

3.2 Estimation Results

The modeling of the inflation rate has been carried out by using monthly US inflation data for the period 01/1954 until 02/2010 taken from Datastream. The time series and its empirical density estimate are plotted in panel (b) and (c) of Fig.2. The specified models and its parameters are given in

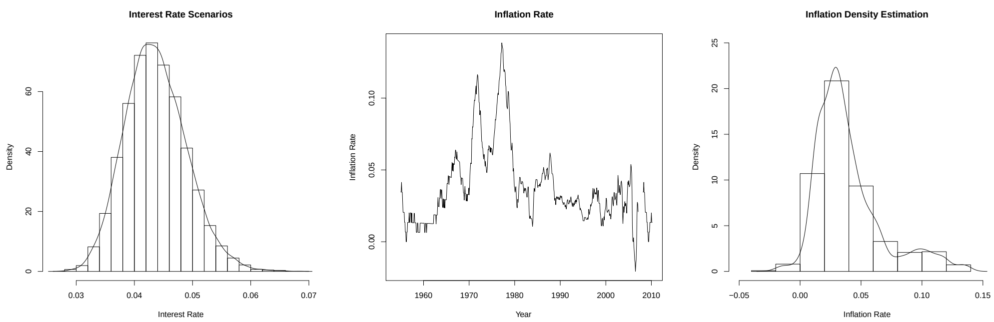

Figure 2: describes the histogram and density estimation of the interest rate scenarios generated by the insurers internal economic scenario generator, the monthly US inflation rate being calculated as the difference of the log consumer price indexes in regard to the respective value from the previous year and the corresponding empirical density estimate.

Tab.1. The *ADF* test (cf. Dickey and Fuller [1979]) as well as the *KPSS* test (cf. Kwiatkowski et al. [1992]) indicate the series to be $I(1)$ which is why henceforward its first difference is utilized. The procedure concerning the simulation of the inflation rates is given as follows. For each of the $I = 15$ models M_i, $i = 1, \ldots, I$, the parameters are estimated. Following forecast values $\hat{I}_{t+h}$ with $h = 1, \ldots, H$ are derived, where the forecast period is chosen to equal $H = 118$. This value accounts for the fact

Equation	i	Model	Parameter										Notation
			p	q	d	p_1	p_2	c	γ	K	L	ψ	
3.3	1	M_1	1	0	0	-	-	-	-	-	-	-	Wilkie, AR(1)
	2	M_2	2^*	2^*	0	-	-	-	-	-	-	-	ARMA(2,2)
	3	M_3	0	0	0.178^{**}	-	-	-	-	-	-	-	ARFIMA(0,d,0)
	4	M_4	2^*	2^*	0.118^{**}	-	-	-	-	-	-	-	ARFIMA(p,d,q)
	5	M_5	0	0	0.083^{**}	-	-	-	-	-	-	-	HSU
3.4	6	M_6	-	-	-	1	1	0.005^{**}	40^{**}	-	-	-	STAR(1,c,γ)
	7	M_7	-	-	-	13^*	13^*	-0.012^{**}	40^{**}	-	-	-	STAR(p,c,γ)
3.5	8	M_8	0	0	0	-	-	-	-	1	0	0	ARCH(1)
	9	M_9	0	0	0	-	-	-	-	4^*	0	0	ARCH(K)
	10	M_{10}	0	0	0	-	-	-	-	1	1	0	GARCH(1,1)
	11	M_{11}	1	0	0	-	-	-	-	1	0	0	ARMA(1,0)-ARCH(1)
	12	M_{12}	1	0	0	-	-	-	-	2^*	0	0	ARMA(1,0)-ARCH(K)
	13	M_{13}	1	1	0	-	-	-	-	1	1	0	ARMA(1,1)-GARCH(1,1)
	14	M_{14}	0	0	0	-	-	-	-	1	0	0.083^{**}	APARCH(1)
	15	M_{15}	0	0	0	-	-	-	-	1	1	0.102^{**}	APARCH(1,1)

Table 1: offers an overview of the specified models. * signifies that the respective lag order has been chosen via information criteria. ** marks estimated values. Note that in M_6 and M_7 γ was respectively estimated to equal 40 signifying that the regime-switch is not carried out smoothly. In fact a threshold autoregressive (TAR) model is specified. The model specifications reported here are the most striking ones regarding its impact on the pension function. We examined a broad range of further specifications which can be reported upon request.

that 3.1 necessitates the 10-year ahead inflation rate while having monthly data up to 02/2010. The forecast values are then given by

$$I_{t+h} = E(I_{t+h}|\Omega_{t+h-1}) + \varepsilon_{t+h}, \quad h = 1, \ldots, H \tag{3.7}$$

where Ω_{t+h-1} is the information set consisting of all relevant information up to and including time $t+h-1$ and ε being an error term. The concrete specifications of the functional form and the error distribution depend on the assumptions of the inflation models being summarized in table 1. For the maximum-likelihood estimation of the models in table 1 we assume gaussian errors.[9]

As we have monthly data the annual inflation rate is given by the mean value over the 12 months of the respective year. This procedure is replicated $N = 10,000$ times for each of the $I = 15$ models yielding the empirical distributions which are summarized in Tab.4 (cf. section A).

[9]We also specified models with t-distributed errors. The results differ quantitatively and not qualitatively from those in tables 1-3.

3.3 Simulation Results

3.3.1 Inflation Models

The results for the resulting inflation distributions are summarized in Tab. 4 (cf. section A). The first striking result marks the fact that the differences of the inflation's distributions mainly focuses on its tails. Whereas the central part of the distributions is surprisingly homogenous the more extreme quantiles and the range differ considerably. This is especially driven by those models belonging to the class of *GARCH* processes (i.e. M_{10}, M_{13} and M_{15}). Although these models forecast rather plausible 10-year ahead inflation rates of approximately 2% in the mean its worst case scenarios of e.g. 170% deflation do not seem to be very realistic.

An explanation for these features can be given by more thoroughly looking at the autocorrelation function of the $GARCH(1,1)$ process. Bollerslev [1986] and Bollerslev [1988] showed that the kth autocorrelation of the squared errors in the $GARCH(1,1)$ process is given by

$$\rho_1 = \alpha_1 + \frac{\alpha_1^2\beta_1}{1-2\alpha_1\beta_1-\beta_1^2} \tag{3.8}$$

$$\rho_k = (\alpha_1+\beta_1)^{k-1}\rho_1 \quad \text{for } k = 2,3,\ldots \tag{3.9}$$

Note that the decay factor of 3.9 is $\alpha_1+\beta_1$. If the sum is close to 1 the autocorrelations will decline only very gradually (although an exponential decline is still given). In our case the sum of the estimated coefficients from the respective *GARCH* models are in each of the three cases very close to 1 i.e. the *GARCH* models feature slowly decaying autocorrelation functions. This leads to the result that draws of exceptionally high error terms during the simulation process 3.7 hardly decline in this model class explaining the extreme scenarios. Note that the fact of the sum of the estimated parameters in $GARCH(1,1)$ models being close to 1 is commonly found in empirical research. E.g. Taylor [1986] estimated $GARCH(1,1)$ models for 40 different financial time series finding in all but six cases that $0.97 \leq \alpha_1+\beta_1 < 1$.

It should furthermore be mentioned that Bollerslev [1986] and Bollerslev [1988] conditioned 3.8-3.9 on the validity of $(\alpha_1+\beta_1)^2+2\alpha_1^2 < 1$ signifying that the kurtosis of ε_t is finite. If however this cannot be maintained, which is the case in our analysis, Ding and Granger [1996] showed that for $\alpha_1+\beta_1 < 1$ and $(\alpha_1+\beta_1)^2+2\alpha_1^2 \geq 1$ the $GARCH(1,1)$ model is still covariance stationary with infinite fourth moment. In this case 3.9 is approximately valid with $\rho_1 \approx \alpha_1+\beta_1/3$. Note also that the *HSU* model (M_5) features a lower mean than the other models. This is due to the fact that we determined $n = 1$ break point via the *LIC* criterion at $t^* = 302$ which corresponds to 02/1979. By looking at Fig.2 it becomes clear that after t^* the trend in the inflation rate is declining what explains the lower mean of M_5 even over 10,000 replications.

The risk measures being defined in section 2 are summarized in Tab.2.

R_{11} for example is calculated as follows. First the $AR(1)$ model

$$I_t = \beta \cdot I_{t-1} + \varepsilon_t \text{ with } \varepsilon \overset{iid}{\sim} N(0,\sigma^2) \tag{3.10}$$

is fitted to the inflation data via maximum-likelihood estimation. The fitted model is then forecasted according to the procedure that is described in section 3.2 where the forecast is carried out under the assumption that the fitted model is indeed the correct model, i.e. that $\hat{\beta} = \beta$ and that

i	Model	π_i	R_{1i}	R_{2i}	R_{3i}	R_{4i}
1	M_1	0.15	1.088	0.816	0.845	0.988
2	M_2	0.15	1.085	0.786	0.807	0.997
3	M_3	0.24	1.096	0.493	0.499	1.016
4	M_4	0.09	1.068	1.408	1.362	1.057
5	M_5	0.15	1.286	0.808	0.814	1.018
6	M_6	0.14	1.098	0.854	0.874	1.000
7	M_7	0.14	1.079	0.813	0.819	1.015
8	M_8	0.12	1.093	0.994	1.029	0.994
9	M_9	0.14	1.116	0.868	1.029	0.868
10	M_{10}	0.21	1.189	0.589	1.029	0.589
11	M_{11}	0.14	1.085	0.849	0.893	0.974
12	M_{12}	0.15	1.088	0.833	0.877	0.975
13	M_{13}	0.26	1.278	0.421	0.134	3.231
14	M_{14}	0.12	1.077	1.055	1.029	1.055
15	M_{15}	0.36	1.568	0.375	1.029	0.375

Table 2: returns measures of market risk (π), estimation risk (R_1), total model risk (R_2), model risk in functional form (R_3) and model risk in distribution (R_4) being defined in section 2 for each of the models in Tab. 1 with $p = 0.99$ and $\alpha = 0.05$.

$\hat{\sigma}^2 = \sigma^2$ in equation 3.10 . Then the 99% VaR of the forecast distribution (corresponding to π_{1,A_1} in definition 1) and the upper bound of the 95% confidence interval for the 99% VaR (corresponding to Π_{1,A_1} in definition 1) is calculated.

As argued is section 2, estimation risk is measured under the assumption that model *i* is the correct model specification. As the correct model is unknown in reality its density function is estimated by an empirical density estimator.[10] Hence feasible versions of the definitions 2-4 are constructed such that under A_1 now the empirical density function is referred to. In this context A_2 now means the possible model specification under model *i*. As the assumed model specification *i* now not necessarily corresponds to the correct one, π_{i,A_2} consequently indicates the 99% VaR under the assumed model specification *i*. For total model risk of model 1 (R_{21}) the numerator in definition 2 is then calculated utilizing the empirical density estimator (corresponding to Π_{A_1} in definition 2).

R_{31} defines model risk in functional form of model 1. Recall that in definition 3, R_{3i} is defined as the upper bound of the confidence interval under the correct model A_1 (i.e. when $a = j$ and $b = k$) divided by the risk measure under a false functional form specification ($a \neq j$ and $b = k$). Without

[10]We used the density estimator implemented in the software *R* which is a nonparametric density estimator using fast Fourier transform.

any functional form specification $I_t = \tilde{\varepsilon}_t$ with $\tilde{\varepsilon} \sim G$ and G denoting the empirical density function of I, i.e. $G = \tilde{f}(I)$. The denominator in definition 3 is now calculated by forecasting the inflation rate based on the model $I_t = \hat{\beta} \cdot I_{t-1} + \tilde{\varepsilon}_t$. Hence the error term is no longer taken to be *iid* normal with variance $\hat{\sigma}^2$ but instead $\tilde{\varepsilon}$ is sampled with replacement from the the empirical density of I. π_{1,A_3} then marks the 99% VaR of the forecast distribution.

Finally concerning R_{41} the definition in remark 2 is utilized. Π_{A_3} is calculated according to the procedure described in the previous paragraph for π_{1,A_3} and π_{1,A_2} is again the 99% VaR of the forecast distribution from $I_t = \hat{\beta} \cdot I_{t-1} + \hat{\varepsilon}_t$ with $\hat{\varepsilon} \overset{iid}{\sim} N(0, \hat{\sigma}^2)$.

By first concentrating on the estimation risk with regard to f_i it becomes clear that for the majority of the models the estimation risk lies somewhere between 5 and 10 percent. The *GARCH* model class again forms an exception with estimation risks up to almost 57% which naturally can be attributed to its near-integratedness and high volatilities.

R_{2i} mirrors total model risk of the underlying, that is specification, estimation and forecasting risk of the respective inflation model. Forecasting risk in this context can be understood as a part of estimation risk as the point estimate of inflation is forecasted. Hence R_{2i} accounts for the risk that the chosen model differs from the true model. Remember that in practice R_{2i} is interpreted as the Basel multiplication factor and is set equal to three apart from some slight possible modifications.[11] Hence utilizing an internal model necessitates to multiply π_i by the factor three when it comes to reporting market risk. Obviously this means that under this regulation there is an incentive to select the model that implies the lowest value of π.

Note that apart from M_4 and M_{14} total model risk is smaller than 1 which is a rather unusual result. A possible explanation is given by the fact that in many popular examples the underlying marks a financial market variable exhibiting the stylized fact of fat-tailedness. This results in the empirical density function having a higher kurtosis than most of the standard parametric models which is why the risk measure of the empirical distribution exceeds the risk measure of the parametric distribution with the resulting multiplication factor exceeding 1. Inflation however is not a monetary but a real variable usually not featuring these stylized facts. Hence in many cases the parametric distribution possesses a much higher kurtosis than its empirical counterpart. Thus our results mirror the following trade-off. Those models implying a low market risk are penalized by a multiplication factor greater than 1 concerning the capital reserves. If a model reports a high market risk it is compensated by a multiplication factor smaller 1.

Focusing on R_{3i} and R_{4i} we can state that model risk in functional form differs much more in between the models than model risk in distribution. This result seems to be rather intuitive as long as the empirical error distribution is close to the normal distribution. Further total model risk is in many cases very close to model risk in functional form indicating that the latter explains a big part of the former. On the other hand model risk in distribution is mostly very close to 1 except for the *GARCH* model class. Note further that there are five models exhibiting the same model risk in functional form of 2.9%. For those models $H(z_t|\theta) = 0$ in 2.1 which means that there are no short-term dependencies in the *DGP*. Hence model risk in functional forms reduces to the estimation error when ε is drawn from the empirical distribution. In other words model risk in distribution exclusively determines total

[11] Precisely the Basel factor additionally covers further effects such as liquidity effects or simplifications of complex positions etc. Hence a comparison should not be understood as being one-to-one.

model risk when there are no short-term dependencies and thus no functional form in the process.

Hence we can conclude that market risk as well as model risk differs substantially between the various econometric models. Whereas theoretically there is a trade-off between market risk and model risk as both depend on the functional form and the distribution of the underlying practically the Basel multiplication factor is fixed which leads to the fact that the model inducing the lowest market risk implies the lowest capital reserves. In this context we do not intent to bother about the (political) question of whether setting $R_2 = 3$ marks a reasonable approach or not. We showed however that if R_2 is interpreted and motivated as a measure for model risk the treatment of holding it constant over different models does not seem to be a plausible approach. The implication of these result in terms of monetary values is outlined next.

3.3.2 Company Model

Once the inflation scenarios have been determined the corresponding pension liabilities are calculated by 3.1 via the calibrated parameters β. As was argued in section 2 the market risk of P if the value-at-risk is chosen as a risk measure is given by $\hat{\pi}_i(p) = \inf\{P \in \mathbb{R} | \int_{-\infty}^{P} \hat{f}_i(P|\hat{\theta};\beta)dL \geq p\}$ with $\hat{f}_i(P) = f_i(x) \cdot |dh(P)/dP|$. 3.1 automatically yields $h(L) = \beta_1^{-1}((L/\beta_0(1+\beta_3 Y)^{-\beta_3})^{1/\beta_1} - 1)$ which is why $|dh(P)/dP| = (\beta_1\beta_2 P)^{-1}(P/\beta_0(1+\beta_3 Y)^{-\beta_3})^{1/\beta_1}$. We may then analyze the impact of the different model risk definitions in terms of monetary values. Since the exact values of $\hat{\pi}$ cannot be reported we again normalize with regard to the benchmark model M_1. Under the Basel accord capital reserves of model M_i in relation to M_1 are given by $CR_{Basel;i} = \hat{\pi}_i \cdot 3/(\hat{\pi}_1 \cdot 3)$ whereas $CR_{New;i} = \hat{\pi}_i \cdot R_{2i}/(\hat{\pi}_1 \cdot R_{21})$ describes the capital reserves under our proposed measure normalized on the first model. As a stylized example capital reserves with respect to market risk are calculated as 8%[12] of the risk weighted assets which are given as a measure for market risk times δ, i.e. $CR = 0.08 \cdot \hat{\pi} \cdot \delta$. Setting $\hat{\pi} = 800$[13] leads to the difference of the two model risk definitions being given as

$$\Delta CR_i = 0.08 \cdot 800(3 - R_{2i}). \tag{3.11}$$

In other words ΔCR_i returns the excess capital reserves in Mio. € under model i when the Basel approach is applied instead of R_{2i}. Tab. 3 mirrors the monetary implications concerning the capital reserves.

It is of no surprise that the results displayed in Tab.2 are rediscovered in the results for CR_{Basel} in Tab.3. As M_4 features the lowest value of the market risk measure π this model exhibits the lowest model risk compensation with respect to a constant multiplication factor. On the other hand M_{15} clearly induces the highest capital reserves due to its high market risk measure. Again the models of the *GARCH* class feature distinct differences in the tails in comparison with the other models. By recapitulating the shape of the pension function (cf. Fig.3) this result should not be surprising. Remember that the *GARCH* inflation scenarios exhibit values in its right tail that are well above 0.4 which is exactly the area where 3.1 increases rapidly. By looking at the pension liabilities' distribution[14] in more detail we can state that as the pension function is leveraged by the inflation scenarios

[12] This value is proposed in the Basel accord (cf. BCBS [1996])

[13] The exact market risk of the insurer's pension liabilities cannot be reported. The amount of 800 Mio.€ is however yields a reasonable approximation.

[14] Note again that we cannot report concrete values of the pension liabilities' distributions due to reasons of concealment.

i	Model	$CR_{Basel;i}$	$CR_{new;i}$	ΔCR_i in Mio.€
1	M_1	1.00	1.00	139.75
2	M_2	1.04	0.96	141.67
3	M_3	1.73	0.60	160.48
4	M_4	0.68	1.72	101.89
5	M_5	1.01	0.99	140.28
6	M_6	0.98	1.05	137.35
7	M_7	1.02	1.00	139.94
8	M_8	0.86	1.22	128.39
9	M_9	0.98	1.06	136.45
10	M_{10}	1.39	0.72	154.25
11	M_{11}	0.99	1.04	137.63
12	M_{12}	1.00	1.02	138.66
13	M_{13}	2.07	0.52	165.08
14	M_{14}	0.82	1.29	124.47
15	M_{15}	2.47	0.46	167.98

Table 3: returns the capital reserves of model i under the Basel accord $(CR_{Basel;i})$ and under R_{2i} in relation to the benchmark model M_1. $CR_{new;i}$ mirrors the difference in million € of of total capital reserves between the Basel approach and our definition of model risk according to 3.11.

the center of the distributions differ slightly more than the scenarios itself. Nevertheless the most striking deviations are once more found in the distributions' tails.

It can further be stated that the (unreported) values for the value-at-risk differ enormously in terms of monetary values. Concretely the discrepancy of the model with the lowest value-at-risk (M_4) and the model with the highest value-at-risk (M_{15}) lies around $5,000$ million €. Of course, one might argue that common sense allows the exclusion of the *GARCH* model class but then still the difference adds up to approximately $2,000$ million € (M_4 vs. M_3). Regarding the expected shortfall the differences in the pension liabilities are even more striking going from $18,000$ million € without exclusion of the *GARCH* model class to $3,000$ million € without consideration of M_{10}, M_{13} and M_{15}. Generally it becomes clear that both a high range and excess kurtosis in the econometric models produces the kurtosis in the pension liabilities' distribution to rise resulting in large values for the value-at-risk and the expected shortfall.

Looking at CR_{new} leads to the opposite implication. Here we observe that the models already entailing high values of π are "compensated" by a low multiplication factor. Intuitively that seems to be a reasonable procedure: If the econometric model induces a rather conservative (i.e. high) value-at-risk the model should be "rewarded" by a smaller multiplication factor if the latter is indeed interpreted as a measure of model risk. Another finding when comparing CR_{Basel} with CR_{New} is that for the former the variation over the models is much stronger. This should also be of no surprise as via consideration of $\tilde{f}(x)$ CR_{New} is a measure relative to the historical data at hand whereas the Basel definition returns a rather absolute parameter. Finally the differences of the two approaches concerning its monetary implications are given by ΔCR. We can see that the difference approximately lies between 100 and 170 million € which makes up about $13-21\%$ of the primary market risk measure. As $R_{2i} < 3 \forall i$ ΔCR is positive for all models meaning that according to our definition of model risk the capital reserves compensation for market risk are smaller than in the Basel approach. This, of course is not a general result but rather caused by the above mentioned situation that the inflations' distributions implied by $M_1 - M_{15}$ are mostly heavier tailed than its empirical counterpart (except for M_4 and M_{14}). If the empirical distribution features heavier tails than the induced model's distribution, ΔCR could very well be negative implying that more capital reserves should be reported in our definition compared to the Basel approach. Hence again we would like to mention that our results should not be interpreted in an absolute way such that our definition of model risk leads to higher or fewer capital reserves. In our opinion however R_2 yields a much better founded multiplication factor if the latter is interpreted as a measure of model risk. With the Basel approach the model returning the lowest market risk measure is "rewarded" in a way that leads to the lowest capital reserves. As the next section shows that marks a very interesting result since M_4 is the model which is indeed chosen by application of an empirical model specification strategy.

3.4 Empirical Model Specification Strategy

The process of finding an appropriate model for the inflation scenarios marks a widely debated task among practitioners. In empirical work a specific model class is often chosen rather ad hoc and a suitable specification procedure is only very rarely carried out. Even if previous work attested a specific model to work very well for the economic variable at interest consulting a different data set might lead to a completely converse implication. That is why we propose a data driven approach concerning the process of model specification. Our strategy consists of at most three steps and is given as follows. At first we try to find the best model in the class of linear time series. Once this model has been found, it is tested for remaining unspecified nonlinearity. If the latter cannot be rejected the best linear model is tested against each of the nonlinear model classes given in Tab.1 being represented by their most general form. In other words we discriminate between the cases of short memory, long memory and spurious long memory. Whilst the decision between short and long memory is the decision between *ARFIMA* and *ARMA*, spurious long memory can be invoked by a nonlinear behavior of the process.

The selection of the most suited linear model might at first sight be thought of as an easy task since almost every conventional time series model is nested in 3.3. I.e. by selecting the lag orders in 3.3 first and estimating the corresponding parameters thereafter one might be able to impose zero restrictions on some of the parameters leading to sub models of the *ARFIMA* class. Hence a general-to-specific modeling procedure equivalent to the Box-Jenkins approach for *ARMA* models might be applied. There are however certain caveats in this argumentation.

Note that there are several ways to estimate the fractional differencing parameter in 3.3 which are based on the periodogram of the process . These include e.g. the *GPH* estimator of Geweke and Porter-Huwak [1983] or the Whittle estimator (cf. Robinson [1995]). The spectrum of a covariance stationary process $\{y_t\}_{t=1}^{T}$ is given as

$$f(\lambda) = |1 - \exp(-i\lambda)|^{-2d} f^*(\lambda), \quad -\pi \leq \lambda \leq \pi, \quad |d| < 0.5 \tag{3.12}$$

with $f^*(\lambda)$ representing the short-term correlation structure of the model and $i = \sqrt{-1}$. In practice 3.12 is approximated by the estimation function

$$\log I(\lambda_k) = c + dX_k + \varepsilon_k, \quad k = 1, \dots, m \tag{3.13}$$

for $m \leq T/2$, $X_k = -2\log(\sin\lambda_k)$, $\lambda_k = 2\pi k/T$ and $I(\lambda) = (2\pi T)^{-1}|\sum_{t=1}^{T} y_t \exp(it\lambda)|^2$ for the sample y_t, $t = 1,\ldots,T$. 3.13 is called the periodogram where c and d can be estimated via linear regression. However, as Hurvich et al. [1998] pointed out, the procedure of estimating d by 3.13 leads to a bias in case there are short-term correlations in the model, i.e. if $f^*(\lambda)$ is not a constant. This induces that if 3.3 contains *ARMA* components no statements about the parameters' significance should be made for estimators based on 3.13.

Hence there are two possibilities for avoiding this shortcoming. Either one applies a different estimator not being based on the periodogram such as the nonparametric estimator proposed by Hurst [1951] or the maximum likelihood estimator of Beran [1995] determining all parameters simultaneously. Or the application of tests discriminating between short and long memory should be carried out. We decided for the second procedure as it is, to our knowledge, not assured that alternative estimators are robust against *ARMA* processes. Concretely we applied two tests in order to discriminate between short and long memory. Firstly we employed the test of Lo [1991] and secondly we applied the test of Davidson and Sibbertsen [2009].

Lo [1991] specifies a modified rescaled range estimator given by

$$\hat{Q}_T = \frac{\max_{0<i\leq T}\{\sum_{t=1}^{i}(y_t-\bar{y})\} - \min_{0<i\leq T}\{\sum_{t=1}^{i}(y_t-\bar{y})\}}{\hat{\sigma}_T} \quad \text{and} \tag{3.14}$$

$$\hat{\sigma}_T = T^{-1}\sum_{t=1}^{T}(y_t-\bar{y})^2 + 2T^{-1}\omega_j(q)\left(\sum_{t=j+1}^{q}(y_t-\bar{y})(y_{t-j}-\bar{y})\right) \tag{3.15}$$

where $\{y\}_{t=1}^{T}$ denotes the process of interest with mean $\bar{y} = T^{-1}\sum_{t=1}^{T} y_t$ and $\omega_j(q) = 1-(j/(q+1))$ for $q < T$. Hence 3.14 can be interpreted as the range of partial sums of deviations of a time series from its mean, rescaled by its standard deviation. Note that 3.15 is the heteroscedasticity and autocorrelation consistent variance estimator with the weights $\omega_j(q)$ being those suggested by Newey and West [1987]. Hence in case the process is short-range dependent $\hat{\sigma}_T$ controls for the autocovariances making 3.14 able to discriminate between short-range and long-range dependence. As $T^{-1/2}\hat{Q}_T$ is asymptotically distributed as the range of a standard brownian bridge under the Null of short-range dependence the latter can be tested against the alternative of long-range dependence.

The bias test of Davidson and Sibbertsen [2009] is based on 3.12 and tests $H_0 : f^*(\lambda) = cons$ vs. $H_1 : f^*(\lambda) \neq cons$. The test statistic is given by

$$TS = \frac{\hat{d}_1 - \hat{d}_2}{SE(\hat{d}_1 - \hat{d}_2)} \tag{3.16}$$

where $\hat{d}_1$ and $\hat{d}_2$ denote alternative estimators of the fractional differencing parameter with $SE(\cdot)$ being a suitable estimation for the difference's standard deviation being derived in Davidson and Sibbertsen [2009]. The authors further proved in their paper that $TS \xrightarrow{d} N(0,1)$ under certain conditions. Choosing $\hat{d}_1$ to be the estimator regressing $I(\lambda_k)$ onto $(X_k, 1)$ in 3.13 whilst $\hat{d}_2$ is derived if $I(\lambda_k)$ is regressed onto $(X_k, 1, h_1(\lambda_k), \ldots, h_{pT}(\lambda_k))$, where $h_j(\lambda_k) = \cos(j\lambda_k)/\sqrt{\pi}$ is the jth order Fourier frequency, one can test the Null of an $ARFIMA(0,d,0)$ process against the alternative of an $ARFIMA(p,d,q)$ process with either $p > 0$ and/or $q > 0$. Note that 3.16 is a simple type of the Hausman [1978] test as under the Null $\hat{d}_1$ is consistent and asymptotically efficient, but biased and inconsistent under the alternative, whereas $\hat{d}_2$ is consistent under both hypotheses.

With a p-value being very close to zero the Lo test clearly rejects the Null of short-range dependence. The p-value of the bias test equals 0.045 indicating that there are *ARMA* components in the process at the 5% level. Having in mind the result of d being actually different from zero we can now be relatively sure that the presence of the fractional differencing parameter is due to long-range dependence and not spuriously caused by *ARMA* components although the latter are indeed present. Thus in the next step we estimated the parameters of the $ARFIMA(p,d,q)$ model simultaneously by the method of Beran [1995] after selecting the lag orders via the Schwarz information criterion leading to the values reported in Tab.4 for M_4.

Once a linear model has been specified and estimated it is tested against remaining nonlinearity. Note that there are several linearity tests in the literature (for an overview cf. Granger and Teräsvirta [1993]). We focused on testing against unspecified remaining nonlinearity as from a practical point of view it is not feasible to carry out different types of tests for every kind of nonlinear model. Thus we applied the popular test of Tsay [1986] performing considerably well in small samples as has been shown in simulation studies (cf. e.g. Tsay [1986] or Pena and Rodriguez [2005]). The test can be described as follows.

At first a linear model (in our case M_4) is fitted to the time series and the residuals of the linear fit $\hat{\varepsilon}_t$ are computed. Secondly $h = M(M+1)/2$ proxy variables, where M stands for the autoregressive order of the process, are defined. The proxy variables are represented by $z_t = vech(Y_t^{'}Y_t)$ where $Y_t = (y_{t-1},\ldots,y_{t-M})$ and $vech(\cdot)$ denotes the column stacking operator using only those elements on or below the main diagonal of each column. Hence z_t consists of several squares and cross products of the series typifying the nonlinearity. Thirdly each of the proxy variables is regressed against Y_t and the h corresponding residuals are denoted by $\hat{u}_t$. Finally the model

$$\hat{\varepsilon}_t = \xi \cdot \hat{u}_t + \iota_t \tag{3.17}$$

where ι is white noise and $\xi = (\xi_1,\ldots,\xi_h)$ denotes a vector of coefficients. 3.17 is estimated by OLS and $H_0 : \xi_1 = \cdots = \xi_h = 0$ vs. $H_1 : \xi_i \neq 0$, for at least one $i = 1,\ldots,h$ is tested consulting a conventional F-test. Under the Null no remaining nonlinearity covered by the proxy variables can be detected. Having utilized the test we do not find remaining nonlinearity in M_4 as the Tsay test reported a p-value of 0.75 for $M = 3$.[15]

4 Conclusion

In this paper we elaborated a definition of model risk as being interpreted as every risk induced by the choice, specification and estimation of a statistical model. We further differentiated model risk into estimation risk, model risk functional form and model risk in distribution. Afterwards we compared our definition of model risk with the standard definition under the Basel accord. As a toy model we looked at the interaction of an econometric model and the corresponding pension liabilities of a large German insurance company as an example of a company model under the focus of induced model risk. For this purpose we modeled the inflation rate with 15 different time series models representing most of the conventional time series models in the literature. We then looked at the impact of model

[15] M was determined by information criteria. Choosing different orders did not alter the test's outcome.

risk on capital reserves with the former functioning as a multiplication factor concerning market risk under the two different definitions of model risk.

The first striking result marks the fact that the different econometric models feature rather decent (between 5 and 10 %) estimation risk. With model risk in distribution also being rather small total model risk is mainly caused by model risk in functional form. By using real insurance data we then determined the corresponding pension liabilities finding that induced model risk differs remarkably depending on the econometric model that is applied. In general the model risk rises if the range and/or the kurtosis of the inflation scenarios increases. Concretely the discrepancy between the models might add up to several million € concerning capital reserves. We found that under the Basel accord the model featuring the lowest market risk measure is ”rewarded” with the lowest resulting capital reserves as the multiplication factor remains constant. In the context of our model risk definition the opposite is true. The model featuring the highest market risk is ”compensated” for its conservative estimation by a low multiplication factor. In our view holding the multiplication factor constant counteracts the motivation of model risk which is to link the capital reserves to the concrete econometric model specification. Our proposed definition overcomes this caveat as the measure is directly depends on the input data at hand. Comparing these two approaches of model risk as to the consequences in terms of capital reserves we find the difference to add up to 100-170 mio. € in our example. Finally we applied a data driven specification strategy in order to specify the inflation model resulting in the model with the lowest market risk and thus the model with the lowest induced capital reserves under the Basel approach being chosen.

Our analysis might be refined in two respects. First, we merely focused on cascade models as economic scenario generators. Here our work might easily be extended via utilization of (structural) multivariate models covering topics such as cointegration or causality. Furthermore consulting an alternative company model being dependent of more than two economic variables offers the analysis of the cascade structure of the Wilkie model in general and possible improvements thereof.

Secondly we did not deal with errors which might occur by leveraging the pension function. I.e. the selection of model points as well as the calibration of the pension function was neglected. Especially the first aspect is worth considering as it is still unclear which scenarios should be selected such that an appropriate fit of the pension function is achieved. Considering that an unrepresentative selection might lead to a bad fit misleading statements concerning the pension liabilities and the induced model risk might be concluded.

A Tables

A.1 Tab.4

	Model														
Statistic	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	M_{11}	M_{12}	M_{13}	M_{14}	M_{15}
Minimum	-0.18	-0.21	-0.33	-0.09	-0.21	-0.16	-0.14	-0.16	-0.25	-1.09	-0.19	-0.20	-1.70	-0.15	-1.24
1%-Quantile	-0.10	-0.11	-0.20	-0.04	-0.13	-0.10	-0.07	-0.08	-0.10	-0.18	-0.10	-0.11	-0.22	-0.08	-0.25
5%-Quantile	-0.07	-0.08	-0.13	-0.03	-0.09	-0.07	-0.04	-0.05	-0.06	-0.08	-0.06	-0.07	-0.11	-0.05	-0.12
1st Quartile	-0.01	-0.02	-0.04	0.00	-0.03	-0.02	0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.02	-0.01	-0.03
Median	0.02	0.02	0.02	0.02	0.01	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Mean	0.02	0.02	0.02	0.02	0.01	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
3rd Quartile	0.06	0.06	0.09	0.04	0.05	0.05	0.07	0.05	0.05	0.05	0.06	0.06	0.06	0.05	0.06
95%-Quantile	0.11	0.12	0.18	0.07	0.11	0.10	0.11	0.09	0.10	0.12	0.11	0.11	0.15	0.09	0.17
99%-Quantile	0.15	0.15	0.24	0.09	0.15	0.14	0.14	0.12	0.14	0.21	0.14	0.15	0.26	0.12	0.36
Maximum	0.22	0.25	0.52	0.12	0.23	0.21	0.22	0.18	0.35	0.84	0.20	0.23	0.74	0.20	2.95
1st Moment	0.02	0.02	0.02	0.02	0.01	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
2nd Moment	0.05	0.06	0.10	0.03	0.06	0.05	0.05	0.04	0.05	0.07	0.05	0.05	0.09	0.04	0.11
3rd Moment	-0.02	-0.03	0.02	0.01	-0.03	-0.05	0.00	-0.03	0.07	-0.33	-0.03	-0.01	-1.4	-0.05	3.42
4th Moment	-0.04	-0.01	0.00	-0.01	-0.02	-0.05	-0.05	0.02	1.14	18.5	0.04	0.15	27.2	0.06	78.53

Table 4: gives some descriptive statistics of the $J = 10,000$ simulated inflation path according to the respective model concerning Tab.1. The 4th moment corresponds to excess kurtosis compared to the normal distribution.

References

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

BCBS. *Amendment to the Basle capital accord to cover market risks*. Basle Committe on Banking Supervision, 1996.

J. Beran. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, No.4:659–672, 1995.

T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31:307–327, 1986.

T. Bollerslev. On the correlation structure for the generalized autoregressive conditional heteroscedastic process. *Journal of Time Series Analysis*, 9:121–131, 1988.

W.A. Brock, S.N. Durlauf, and K.D. West. Policy analysis in uncertain economic environments. *Brookings Papers on Economic Activity*, 1:235–332, 2003.

K.S. Chan and H. Tong. On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7:179–190, 1986.

R. Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16:519–547, 2006.

M. Crouhy, D. Galai, and R. Mark. Model risk. *Journal of Financial Engineering*, 7:267–288, 1998.

J. Davidson and P. Sibbertsen. Test of bias in log-periodogram regression. *Economics Letters*, 102: 83–86, 2009.

D.A. Dickey and W.A. Fuller. Distribution of the estimator for autoregressive time series with unit root. *Journal of the American Statistical Association*, 74:427–431, 1979.

F. Diebold and A. Inoue. Long memory and regime switching. *Journal of Econometrics*, 105:131–159, 2001.

Z. Ding and C.W.J. Granger. Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics*, 73:185–215, 1996.

Z. Ding, C.W.J. Granger, and R.F. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106, 1993.

R.F. Engle and A.D. Smith. Stochastic permanent breaks. *Review of Economics and Statistics*, 81: 553–574, 1999.

J. Geweke and S. Porter-Huwak. The estimation and application of long-memory time series models. *Journal of Time Series Analysis*, 4:221–238, 1983.

C.W.J. Granger and T. Teräsvirta. *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press, 1993.

J.A. Hausman. Specification tests in econometrics. *Econometrica*, 46:1251–1272, 1978.

C.-C Hsu. Long memory or structural changes: An empirical examination on inflation rates. *Economics Letters*, 88:289–294, 2005.

H.E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799, 1951.

C.M. Hurvich, R. Deo, and J. Brodsky. The mean squared error of geweke and porter-hudak's estimator of a long memory time series. *Journal of Time Series Analysis*, 19:19–46, 1998.

J. Kerkhof, B. Melenberg, and H. Schumacher. Model risk and capital reserves. *Journal of Banking and Finance*, 34:267–279, 2010.

D. Kwiatkowski, P.C. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stionarity against the alternative of a unit root. *Journal of Econometrics*, 54:159–178, 1992.

M. Lavielle and E. Moulines. Least squares estimation of an unknown number of shifts in a time series. *Journal of Time Series Analysis*, 21:33–59, 2000.

A.W. Lo. Long-term memory in stock market prices. *Econometrica*, 59, No.5:1279–1313, 1991.

W. Newey and K. West. A simple positive definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–705, 1987.

D. Pena and J. Rodriguez. Detecting nonlinearity in time series by model selection criteria. *International Journal of Forecasting*, 21:731–748, 2005.

P. Robinson. Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, 23:1630–1661, 1995.

J.V. Rosenberg and T. Schuermann. A general approach to integrated rsik management with skewed fat-tailed risks. *Journal of Financial Economics*, 79:569–614, 2006.

G. Stahl. Three cheers. *Risk Magazine*, 10:67–69, 1997.

S. Taylor. *Modeling financial time series*. Wiley, New York, NY, 1986.

T. Teräsvirta. Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89:208–218, 1994.

R.S. Tsay. Nonlinearity tests for time series. *Biometrika*, 73:461–466, 1986.

H.G. Tucker. A generalization of the glivenko-cantelli theorem. *The Annals of Mathematical Statistics*, 30(3):828–830, 1959.

A. Wilkie. More on a stochastic asset model for actuarial use. *British Actuarial Journal 1*, V:777–964, 1995.