MODEL RISK AND REGULATORY CAPITAL

By Jeroen Kerkhof, Bertrand Melenberg and
Hans Schumacher

April 2002

Discussion paper

Model Risk and Regulatory Capital

JEROEN KERKHOF[*], BERTRAND MELENBERG[†], and HANS SCHUMACHER[‡]

Tilburg University

PO Box 90153

5000 LE Tilburg

The Netherlands

April 2, 2002

[*]Department of Econometrics and Operations Research, and CentER, Tilburg University, and Product Development Group, ABN◇AMRO Bank, Amsterdam. E-mail: F.L.J.Kerkhof@kub.nl, Phone: +31-13-4662134, Fax: +31-13-4663280.

[†]Department of Econometrics and Operations Research, Department of Finance, and CentER, Tilburg University. E-mail: B.Melenberg@kub.nl.

[‡]Department of Econometrics and Operations Research, and CentER, Tilburg University. E-mail: J.M.Schumacher@kub.nl.

Model Risk and Regulatory Capital

ABSTRACT

In this paper we propose a general framework for quantification of model risk. This framework allows one to allocate regulatory capital to positions in a given market depending on the extent to which this market can be reliably modeled. Our approach is based on computing worst-case risk measures over sets of models that are in some appropriate sense close to a nominal model. The method is general in the sense that it can be applied with any of the usual risk measures such as Value-at-Risk and Tail Conditional Expectation. Inasfar as risk measures can also be used as pricing tools or as determinants of margin requirements, the paper provides a quantification of model risk in these settings as well. We present applications both to stock portfolios and to derivative products: we find that, for usual specifications, misspecification risk is much more important than estimation risk.

JEL codes: G12, G18

Key words: Capital requirements, (Coherent) Risk management, Derivative pricing models.

I. Introduction

Due to the growing complexity of financial markets, financial institutions rely more and more on the use of models to assess the risks to which they are exposed. The accuracy of these risk assessments depends crucially on the extent to which a market can be reliably modeled. Choosing an appropriate model to compute market risk measures is an important and difficult task. It is a widespread feeling among both academics and practitioners that, although some models do a better job than others, the search for one ultimate model is futile. An approach that takes the limitations of our knowledge into account is to develop models—depending on the application—that capture the most important aspects of a particular market, and to somehow control for the fact that the assessment of risk is based on a possibly misspecified model (see Derman (1996)).

The hazard of working with a potentially misspecified model is called *model risk*. Currently no explicit capital requirements are set by the regulators in connection with model risk. This done indirectly using the so called multiplication factors. However, the Basle Committee has indicated that it plans to expand the current capital adequacy framework to improve the charting of risks to which financial institutions are exposed (see Basle Committee on Banking Supervision (1999)). In particular, the Committee intends to set capital requirements for operational risk, which consists for an important part of model risk. Just as the 1996 Amendment of the Basle Committee stimulated financial institutions to refine their market risk models, banks are likely to make more detailed assessments of model risk after incorporation of model risk regulation in the Basle Accord. As part of their internal risk management systems, most large financial institutions already set aside reserves for model risk (the so-called *model reserves*). This means that booking of certain profits on trades is postponed if it is felt that these profits are sensitive to the used model.

The aim of this paper is to provide a quantitative basis for the incorporation of model risk in regulatory capital requirements. The same framework may also be used for the computation of model reserves in the context of internal risk management procedures within financial institutions; in addition, the method may be used in margin setting

by clearing house exchanges, or as a pricing tool. To extend the current practice of computing market risk measures on the basis of some given ("nominal") model, we determine a set of plausible alternative models. In recognition of the fact that each of these models is a (reasonable) candidate for representing reality, we propose to compute a worst-case market risk measure over our set of alternative models. Model risk is then defined as the difference between this measure and the market risk measure computed from the nominal model. We distinguish between model risk due to estimation error and model risk due to misspecification.

Previous studies on model risk in interest rate markets have focused on the risk of using incorrect parameter values in a parametric setting (see, for example, Gibson, Lhabitant, Pistre, and Talay (1999), Talay and Ziyu (2000), Bossy, Gibson, Lhabitant, Pistre, and Talay (2000), and Bossy, Gibson, Lhabitant, Pistre, Talay, and Ziyu (2000)). However, our study suggests that the leading factor in model risk is often misspecification rather than parameter estimation error. Hull and Suo (2001) investigate the model risk associated with the calculation of prices and deltas for illiquid exotic options based on an implied-volatility model that is calibrated using current prices of liquid products. Their paper clearly demonstrates the presence of model risk in a number of situations. In this paper, we assess risk in derivative products on the basis of the total hedging error rather than the error in computing Greeks, and we propose a quantitative measure of model risk that could be used, for instance, in the determination of model reserves. Steps towards the quantification of model risk for derivative contracts have been taken by Green and Figlewski (1999), who show that the risk of trading derivative securities can be decreased substantially by delta hedging. We follow this line of thought below by considering the risk of derivative products in combination with a given hedging strategy. The proposed methodology encompasses the methodology proposed by Hull and Suo (2001). Furthermore, robustness issues as treated by El Karoui, Jeanblanc-Picqué, and Shreve (1998) fit into the proposed setup.

One major area where financial models play an important role is the risk management of the portfolios of financial institutions. We discuss value-at-risk and tail conditional

expectation for a simple model to illustrate our model risk measurement tools. The results can be interpreted in terms of a multiplication factor that should be applied to account for model risk in a given market. Our results for this simple model indicate that a significant part of the multiplication factor of three to four used by the Basle Committee can be explained in this way when computing the 99% value-at-risk at a 95% confidence level. We find that the model risk due to misspecification is much larger than the model risk due to estimation error.

Another area which relies heavily on financial models is constituted by derivatives trading. As already mentioned, Green and Figlewski (1999) found that risks in derivative contracts can be reduced considerably by delta hedging. To evaluate derivative pricing models it is, therefore, important to take the hedging strategy into account as well, since financial institutions use the possibility of hedging extensively. A natural and intuitive approach is to view the cost of hedging of a derivative using a particular hedging strategy, instead of the final payoff as the risk of a derivative. Since most derivative pricing models are continuous-time models while in practice it is only possible to hedge in discrete time, derivative contracts cannot be hedged perfectly and so they are subject to market risk. The assessment of this market risk is model-dependent, and we define model risk as the difference between the worst-case market risk measure and the nominal market risk measure. We illustrate our approach for the Black-Scholes family of option pricing models. Our results indicate again that model risk due to misspecification is much larger than the model risk due to estimation error.

The remainder of the paper is structured as follows. In the next section, we give an overview of market risk measurement. We discuss some of the popular risk measures with some emphasis on coherent market risk measurement which fits neatly with the model risk measurement method proposed in Section III. In Section III we propose a general framework for incorporation of model risk. This is based on a type of worst-case analysis. A decomposition of model risk in a parametric and a nonparametric part is proposed. Section IV provides an application to portfolio risk management. We discuss the value-at-risk and the tail conditional expectation approach. In Section V we apply

our methodology to derivative securities. Finally, Section VI concludes.

II. Market Risk Measurement

By *market risk* we understand the risk caused by random fluctuations in future asset prices. For each given position, the most basic question that a risk manager must be able to answer is whether or not the risk associated to this position is acceptable. This qualitative decision is often based on the computation of a "risk measure" which in some way represents the distance to (un)acceptability. Such a risk measure may, for instance, be arrived at as follows. Since, in the context of finance, risk is usually measured in terms of the distribution of only one variable (profit/loss), an unacceptable position can be made acceptable if enough of a suitable "sweetener" is added. The amount of sweetener that has to be added to make a given position just acceptable is a natural measure of the distance to acceptability.

A. Use of market risk measures

Market risk measures may be used for a number of different purposes.

1. Regulatory capital requirements for securities firms and banks are computed on the basis of risk measures. Specifically, the value-at-risk method has been adopted by the Basle Committee on Banking Supervision (1996).

2. Some banks set reserves for trading desks as part of their internal risk management procedures. The size of the reserve is coupled to some measure of the riskiness of the positions taken by the desk.

3. Exchanges need to guarantee the promises to all parties involved in a contract. To guarantee these promises they use clearing margins for their members. For example, the Chicago Mercantile Exchange (CME) and many other exchanges use SPAN to determine the clearing margins. For more detailed information, see Artzner, Delbaen, Eber, and Heath (1999) and SPAN (1995).

6

4. Market risk measures may also be used as pricing tools, since they can be used to compare different risks so that good deals can be identified. This point of view is elaborated for instance in Cochrane and Saa-Requejo (2000) and Jaschke and Küchler (2001).

B. Notation and definitions

Since in this paper we are interested in model risk, we will be working with classes of models rather than with a single model. It is not always convenient to use the same probability space for each of these models. Therefore, we start by a formal description of a setting that allows the use of multiple probability spaces.

Definition 1 A *model* is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

One could imagine more elaborate probabilistic settings; in particular, a filtration might be assumed given. However, the above notion will be sufficient for the purposes of this paper. For any model m, let $\mathcal{R}(m)$ denote the space of equivalence classes of measurable real-valued functions on $(\Omega, \mathcal{F})$.

Definition 2 Let a model m be given. A *risk* defined on m is an element of $\mathcal{R}(m)$.

This definition, in which a "risk" is a random variable defined on a given probability space, follows the terminology of Artzner, Delbaen, Eber, and Heath (1999) and Delbaen (2000). We introduce a similar concept for model classes rather than for individual models.

Definition 3 Let $\mathcal{M}$ be a class of models. A *product* defined on $\mathcal{M}$ is a mapping that assigns to each model $m \in \mathcal{M}$ a risk defined on m. The set of all products defined on $\mathcal{M}$ is denoted by $\mathcal{X}(\mathcal{M})$.

The risk induced by a product Π on a model m will be denoted by Π_m. Since $\mathcal{R}(m)$ is a vector space, the set of products $\mathcal{X}(\mathcal{M})$ has the structure of a vector space as well. For instance, if Π_1 and Π_2 are products defined on the same class of models $\mathcal{M}$, then

7

$\Pi_1 + \Pi_2$ is the product that associates to a model m in $\mathcal{M}$ the risk $(\Pi_1)_m + (\Pi_2)_m$. Similarly, we can also define products relative to a reference product (if the reference product is nonzero), and we have a partial ordering on products.

We now proceed to risk measures, again starting with the definition for an individual model.

Definition 4 Let a model m be given. A *risk measure* defined on m is a map from $\mathcal{R}(m)$ to $\mathbb{R} \cup \{\infty\}$.[1]

Definition 5 Let a class of models $\mathcal{M}$ be given. A *risk measurement method* defined on $\mathcal{M}$ is a mapping that assigns to each model $m \in \mathcal{M}$ a risk measure defined on m.

Risk measures can be used to separate "acceptable" from "unacceptable" risks in the following way.

Definition 6 Let a model m be given, and let ρ be a risk measure defined on m. The acceptance set associated with ρ is the set

$$\mathcal{A}_\rho = \{X \in \mathcal{R}(m) \mid \rho(X) \leq 0\}. \tag{1}$$

So far we did not discuss specific properties for risk measures and related notions that would justify the nomenclature. We come to this in the next section.

C. Popular risk measurement methods and their properties

Most risk measures used in practice can be viewed as risk measurement methods in the formal sense of the previous section. Due to its prominent role in the amendment of 1996 by the Basle Committee, the value-at-risk approach is currently the most popular method used in risk measurement (see, for example, Duffie and Pan (1997), Basle Committee on Banking Supervision (1996), Dowd (1998), and RISK Magazine (1996)). A formal description of VaR may be given as follows.

---

[1] Including ∞ allows risks to be defined on more general probability spaces, see Delbaen (2000).

Definition 7 (Value at Risk (VaR)) Let a model class $\mathcal{M}$ be given. The *value-at-risk* method with *reference asset* $N \in \mathcal{X}(\mathcal{M})$ and *level* $p \in (0,1)$ assigns to a model $m = (\Omega, \mathcal{F}, \mathbb{P}) \in \mathcal{M}$ the risk measure VaR_m given by

$$\mathrm{VaR}_m : \mathcal{R}(m) \ni X \mapsto -\inf\{q \in I\!R : \mathbb{P}(X/N_m \leq q) \geq p\} \in I\!R \cup \{\infty\}. \qquad (2)$$

We now list a number of properties that risk measures and risk measurement methods may satisfy. Again, we start with individual models. So, let a model m be given, and let ρ be a risk measure defined on m. Since m will be fixed for the moment, we write $\mathcal{R}(m)$ simply as $\mathcal{R}$. Some properties of interest will be stated as axioms. In the first axiom we also assume that a reference risk $N \in \mathcal{R}$ has been given.

Axiom 1 (Translation invariance) For all $X \in \mathcal{R}$ and $\tau \in \mathbb{R}$, we have $\rho(X + \tau N) = \rho(X) - \tau$.

Adding (subtracting) an initial investment of size τ in the reference asset N decreases (increases) the risk measure ρ by τ. Therefore, τ can be interpreted as the amount of the sweetener for the risk X to become more acceptable (or less, in case τ is negative).

Axiom 2 (Monotonicity) For all X and $Y \in \mathcal{R}$ with $X \leq Y$, we have $\rho(X) \geq \rho(Y)$.

It seems natural to assign a higher value to risks that always have a lower payoff. Note that the axiom of monotonicity rules out the commonly used mean-variance measure $\rho(X) = -\mathbb{E}_{\mathbb{P}}(X) + \gamma \mathbb{V}\mathrm{ar}_{\mathbb{P}}(X)$, where γ is a risk aversion parameter. The VaR measure, on the other hand, is monotonic.

Axiom 3 (Positive homogeneity) For all $X \in \mathcal{R}$ and $\lambda \geq 0$, $\rho(\lambda X) = \lambda \rho(X)$.

Again, this axiom is satisfied by VaR. The homogeneity axiom may be considered reasonable as a local approximation, or when size effects (due, for instance, to liquidity risk or to regulatory constraints) are taken into account in the future net worth of a position.

Axiom 4 (Subadditivity) For all X and $Y \in \mathcal{R}$, we have $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

If the risk measure ρ satisfies the subadditivity property, the risk manager/supervisor is sure that the sum of two separate risks X and Y can be estimated conservatively by the sum of the risk measures of the separate risks. If a risk measure does not satisfy the subadditivity property, a risk might be disguised by splitting it up. The VaR measure does not satisfy the subadditivity property (see Artzner, Delbaen, Eber, and Heath (1999) for a counterexample).

The above axioms can be transferred to risk measurement methods in a straightforward way. We shall say that a risk measurement method RMM defined on a model class $\mathcal{M}$ satisfies Axiom i $(i = 1, \ldots, 4)$ if for each $m \in \mathcal{M}$ the risk measure RMM_m on m satisfies Axiom i with $\mathcal{R} = \mathcal{R}(m)$. In the case of the translation invariance axiom, it is assumed that a reference product in $\mathcal{X}(\mathcal{M})$ is given.

The fact that VaR does not satisfy the subadditivity property is often seen as a disadvantage of this risk measurement method; see Artzner, Delbaen, Eber, and Heath (1999) for a more extensive discussion. Alternative risk measures have been proposed that do satisfy the desirable subadditivity property. Artzner, Delbaen, Eber, and Heath (1997) introduced the notion of coherent risk measures. Their ideas were formalized in Artzner, Delbaen, Eber, and Heath (1999), Artzner (1999), and Delbaen (2000).

Definition 8 A *coherent risk measure* is a risk measure that satisfies the axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity.

The definition can immediately be extended to produce the notion of a *coherent risk measurement method*.

The four axioms still allow many measurement methods, so even when one decides to use a coherent measure one needs further considerations to arrive at a specific method. An example of a coherent risk measurement method is the worst conditional expectation (WCE). Contrary to VaR, this measure takes the size of losses under the VaR limit into account. Therefore, it is not possible to increase the expected return of a portfolio under WCE restrictions by taking extremely risky bets with a very low probability but a very high loss.

Definition 9 (Worst Conditional Expectation (WCE)) Let a model class $\mathcal{M}$ be given. The *worst conditional expectation* method with *reference product* $N \in \mathcal{X}(\mathcal{M})$ and *level* $p \in (0,1)$ assigns to a model $m = (\Omega, \mathcal{F}, \mathbb{P}) \in \mathcal{M}$ the risk measure WCE_m given by

$$\text{WCE}_m : \mathcal{R}(m) \ni X \mapsto -\inf_{A \in \mathcal{F}, \, \mathbb{P}(A) > p} \mathbb{E}_{\mathbb{P}} [X/N_m \mid A] \in I\!R \cup \{\infty\}. \qquad (3)$$

It is a straightforward exercise to show that WCE satisfies the axioms of translation invariance, monotonicity, positive homogeneity, and subadditivity. Though WCE has nice theoretical implications, it is difficult to compute in practice. However, it can be shown that WCE equals the tail conditional expectation (TCE) in all practically relevant cases (see proposition 5.3 of Artzner, Delbaen, Eber, and Heath (1999)). TCE is defined as follows.

Definition 10 (Tail Conditional Expectation (TCE)) Let a model class $\mathcal{M}$ be given. The *tail conditional expectation* method with *reference product* $N \in \mathcal{X}(\mathcal{M})$ and *level* $p \in (0,1)$ assigns to a model $m = (\Omega, \mathcal{F}, \mathbb{P}) \in \mathcal{M}$ the risk measure TCE_m given by

$$\text{TCE}_m : \mathcal{R}(m) \ni X \mapsto -\mathbb{E}_{\mathbb{P}} [X/N_m \mid X/N_m \leq -\text{VaR}_m(X)] \in I\!R \cup \{\infty\} \qquad (4)$$

where VaR is taken with reference product R and at level p.

III. Model risk

Market risk measures are typically based on a class of scenarios together with a base probability measure; both items are provided by a model m. At a higher level, however, there is uncertainty about which model to use. A financial institution's perception of market risk can deviate substantially from the true market risk due to the fact that the actual dynamics are insufficiently represented by the model dynamics. Due to the use of an incorrect model, the financial institution may accept risks that it would find unacceptable in case it would know the actual dynamics. The risk associated to the mismatch between model dynamics and actual dynamics is called *model risk*.

We now want to find a quantitative measure for model risk. Since the true dynamics are unknown, it makes sense to form a set of alternative dynamics $\mathcal{K}$ (containing a nominal model m) which is likely to contain the true dynamics. A natural candidate for a model risk measure is the difference between the worst-case risk measure among all models in the neighborhood $\mathcal{K}$ and the risk measure under the dynamics of the nominal model m. If the market risk measurement method is translation invariant, the difference between these two quantities gives the extra position in the reference product which has to be added to the market risk measure of the nominal model to make the risk acceptable even under the worst case dynamics. In the next section, this intuition is formalized.

A. Measuring model risk

Suppose that the financial institution uses a risk measurement method RMM to assess the acceptability of a product (portfolio) Π. In model m, the risk of the product Π is computed as $\mathrm{RMM}_m(\Pi_m)$. To take into account model uncertainty, we take a set of alternative dynamics $\mathcal{K}$ around m and compute the *worst-case market risk measure* (with respect to $\mathcal{K}$), which is given by $\sup_{k \in \mathcal{K}} \mathrm{RMM}_k(\Pi_k)$. Model risk may now be quantified as follows.

Definition 11 (Model risk measure)[2] Let $\mathcal{M}$ be a class of models, let m be a model in $\mathcal{M}$, and let $\mathcal{K}$ be a subset of $\mathcal{M}$ containing m. Furthermore, let Π be a product defined on $\mathcal{M}$ and let RMM be a risk measurement method for $\mathcal{M}$. The *model risk* associated to the method RMM of product Π, with respect to the *nominal model* m and the *tolerance set* $\mathcal{K}$, is given by

$$\phi_{\mathrm{RMM}}(\Pi, m, \mathcal{K}) = \sup_{k \in \mathcal{K}} \mathrm{RMM}_k(\Pi_k) - \mathrm{RMM}_m(\Pi_m). \tag{5}$$

Artzner, Delbaen, Eber, and Heath (1999) and Delbaen (2000) use one particular model m to compute the market risk measure $\mathrm{RMM}_m(\Pi_m)$. With the definition above we extend their risk measurement framework by including model risk. The amount

---

[2]The case where $\mathrm{RMM}_m(\Pi_m) = \infty$ is uninteresting since the financial institution will never accept the product Π in its portfolio.

$\phi_{\mathrm{RMM}}(\Pi, m, \mathcal{K})$ can be thought of as a model reserve that should be held to cover the worst-case dynamics of $\mathcal{K}$. Consider, for example, value-at-risk. From empirical data we can determine whether the VaR limit given by a nominal model is exceeded as often as predicted or more often. If the model is accurate in predicting the VaR limit we would like to set a small model reserve. On the other hand, we want to set a large model reserve in case the model does a poor job predicting the VaR limit. Adding the model reserve $\phi_{\mathrm{RMM}}(\Pi, m, \mathcal{K})$ to the nominal market risk measure $\mathrm{RMM}_m(\Pi_m)$ gives a total risk measure equal to $\sup_{k \in \mathcal{K}} \mathrm{RMM}_k(\Pi_k)$. In appendix B we illustrate the procedure for coherent risk measures, in particular, the WCE and SPAN. The size of the model reserve (and thereby the total risk measure) is controlled by the size of $\mathcal{K}$. In the next section, we discuss the determination of $\mathcal{K}$ and the dependence of the model reserve on model accuracy in more detail.

The model risk measure that we have defined may have some desirable properties depending on the market risk measurement method from which it has been derived.

Theorem 1 (Invariance) Let RMM be a risk measurement method that is translation invariant with respect to a reference product N. Then the model risk measure associated to RMM is translation invariant in the sense that

$$\phi_{\mathrm{RMM}}(\Pi + \tau N, m, \mathcal{K}) = \phi_{\mathrm{RMM}}(\Pi, m, \mathcal{K}) \tag{6}$$

for all $\tau \in \mathbb{R}$.

Proof. Take $\tau \in \mathbb{R}$. We have

$$
\begin{aligned}
\phi_{\mathrm{RMM}}(\Pi + \tau N, m, \mathcal{K}) &= \sup_{k \in \mathcal{K}} \mathrm{RMM}_k\left(\Pi_k + \tau N_k\right) - \mathrm{RMM}_m\left(\Pi_m + \tau N_m\right) \\
&= \sup_{k \in \mathcal{K}} \mathrm{RMM}_k\left(\Pi_k\right) - \tau - \mathrm{RMM}_m\left(\Pi_m\right) + \tau \\
&= \phi_{\mathrm{RMM}}(\Pi, m, \mathcal{K}).
\end{aligned}
$$

∎

The addition of a constant payoff should not alter the model risk, since it is model independent. Another way to look at it is that the constant payoff can be fully hedged by a position in the reference product.

Theorem 2 (Positive homogeneity) Let RMM be a positively homogeneous risk measurement method. Then the model risk measure associated to RMM is positively homogeneous as well, i.e.,

$$\phi_{\text{RMM}}(\lambda\Pi, m, \mathcal{K}) = \lambda\phi_{\text{RMM}}(\Pi, m, \mathcal{K}) \tag{7}$$

for all $\lambda \geq 0$.

Proof. Take $\lambda \geq 0$. We have

$$
\begin{aligned}
\phi_{\text{RMM}}(\lambda\Pi, m, \mathcal{K}) &= \sup_{k\in\mathcal{K}} \text{RMM}_k(\lambda\Pi_k) - \text{RMM}_m(\lambda\Pi_m) \\
&= \lambda\left(\sup_{k\in\mathcal{K}} \text{RMM}_k(\Pi_k) - \text{RMM}_m(\Pi_m)\right) \\
&= \lambda\phi_{\text{RMM}}(\Pi, m, \mathcal{K}).
\end{aligned}
$$

∎

If market risk is measured proportionally to position size, then the same property holds for model risk. This seems a reasonable characteristic.

Our model risk measure does not, in general, satisfy the subadditivity property. However, if the basic market risk measurement method is subadditive, this property does hold for what might be called *total market risk*, viz. the sum of nominal market risk and model risk. This is immediately seen from the fact that total market risk is given by the formula $\sup_{k\in\mathcal{K}} \text{RMM}_k(\Pi_k)$, and from the general fact that $\sup_i(a_i + b_i) \leq \sup_i(a_i) + \sup_i(b_i)$. As noted above, the reason why market risk measures are often required to be subadditive is to prevent companies, trading desks, etc. from covering up large risks by splitting them into separate positions that do satisfy the risk criteria. If total market risk is reported, then subadditivity of this risk measure is sufficient for this.

14

We choose a worst-case approach to quantify model risk. An alternative would be a Bayesian approach, in which the model risk measure is a weighted average of risk measures according to some prior. Depending on its risk attitude, the financial institution can give more weight to unfavorable dynamics. However, the choice of a prior is difficult and arbitrary. In a worst-case approach, one only needs to specify the tolerance set $\mathcal{K}$; this may be seen as an acknowledgment of the restrictions of statistical modeling in the face of limited data and limited understanding of the true dynamics.

B. Decomposition of Model Risk

The exposition given in the previous section was rather general. We did not specify a model m or a set of alternative models $\mathcal{K}$. In this section we discuss some possible choices for the set of alternative dynamics $\mathcal{K}$. In practice, one starts with a (usually parametric) model class, say $\mathcal{M}(\Theta) \equiv \{(\Omega, \mathcal{F}, \mathbb{P}_\theta) : \theta \in \Theta\} \subset \mathcal{M}$. Using an estimation or calibration procedure, a particular element $m(\hat{\theta})$ is chosen from $\mathcal{M}(\Theta)$. Even if the actual dynamics, say m_0, belong to the parametric model class $\mathcal{M}(\Theta)$, that is $m_0 = m(\theta_0)$ for some $\theta_0 \in \Theta$, the financial institution faces the risk of selecting the wrong element $m(\hat{\theta})$. This risk is called *model risk due to estimation error*. To define a neighborhood of plausible values around $m(\hat{\theta})$, one typically uses confidence regions. Specifically, we can place a confidence region around the estimator $\hat{\theta}$ for θ_0 to define some neighborhood around $m(\hat{\theta})$. Depending on a chosen level α we take a $(1-\alpha)\%$ confidence region around $\hat{\theta}$.[3] In this way we arrive at a set of alternative models of the following form:

$$\mathcal{K}(\alpha) = \left\{ m(\theta) \in \mathcal{M}(\Theta) : \theta \in CI_{1-\alpha}(\hat{\theta}) \right\} \tag{8}$$

In situations where one is interested in a specific market risk measurement method RMM and a specific product Π, an alternative approach which focuses more directly on the given situation is to use the set $\mathcal{K}$ defined by

$$\mathcal{K}(\alpha) = \left\{ m(\theta) \in \mathcal{M}(\Theta) : \text{RMM}_{m(\theta)}\left(\Pi_{m(\theta)}\right) \in CI_{1-\alpha}\left(\text{RMM}_{m(\hat{\theta})}\left(\Pi_{m(\hat{\theta})}\right)\right) \right\}. \tag{9}$$

---

[3] $CI_{1-\alpha}(\hat{\theta})$ denotes the $(1-\alpha)\%$-confidence interval for θ_0.

We define *model risk due to estimation error*, or simply *estimation risk*, as the model risk that is obtained from a tolerance set derived from confidence regions in within the model class.

Now let us consider the situation where the actual dynamics may not belong to $\mathcal{M}(\Theta)$. The risks that we are considering are real-valued random variables and so a natural idea is to work on the basis of the associated distribution functions. Suppose that a cumulative distribution function $\hat{F}(x)$ has been obtained by some nonparametric estimation method. This allows us to define a tolerance set $\mathcal{K}$ depending on confidence level α in the following way:

$$\mathcal{K}(\alpha) = CI_{1-\alpha}(\hat{m}) := \left\{ (\Omega, \mathcal{F}, \mathbb{P}) : F(x) = \mathbb{P}((-\infty, x]) \in \left[\hat{F}(x) \pm \frac{k_{\alpha/2}}{\sqrt{n}} \right] \ \ \forall x \in \mathbb{R} \right\},$$

where $k_{\alpha/2}$ is the critical value of the Kolmogorov-Smirnov statistic.[4] As above, one may also define tolerance sets that are more specifically tied to a given risk measurement method and a given product. Along this line, one may estimate $\mathrm{RMM}_m(\Pi_m)$ first and define a tolerance set based on a confidence region $CI_{1-\alpha}$ for the estimate

$$\mathcal{K}(\alpha) = \{m : \mathrm{RMM}_m(\Pi_m) \in CI_{1-\alpha}\}.$$

In general, we can determine tolerance sets that are restricted or are not restricted to a model class (unrestricted, in the sequel). As above, one may define model risk due to estimation error as the model risk restricted to the model class $\mathcal{M}(\Theta)$. The amount that has to be added to arrive at the model risk determined from the unrestricted method may be termed *model risk due to misspecification* or simply *misspecification risk*. In other words, if $\mathcal{K}_r$ is the restricted tolerance set and $\mathcal{K}_u$ is the unrestricted one, then we define the misspecification risk for a given product Π as

$$\phi_{\mathrm{RMM}}(\Pi, \mathcal{K}_r, \mathcal{K}_u) = \sup_{k \in \mathcal{K}_u} \mathrm{RMM}_k(\Pi_k) - \sup_{k \in \mathcal{K}_r} \mathrm{RMM}_k(\Pi_k). \tag{10}$$

---

[4]Alternatively, uniform confidence bounds around a non-parametric distribution may be obtained from the Cramér-von Mises statistic or the Kuiper statistic (see, for example, Shorack and Wellner (1986)).

However, the quantity defined above may in some cases be less than zero, whereas we would prefer to define misspecification risk in such a way that it is always nonnegative. To achieve this with the above definition, we have to make sure that the set $\mathcal{K}_r$ is nested in $\mathcal{K}_u$. One way to ensure nesting is to form convex combinations. Note that, in a context in which we are concerned with a specific product, it is reasonable to identify models with the cumulative distribution functions induced by the given product, and in this way it is indeed possible to consider convex combinations of models. The nesting property can then be guaranteed by replacing the set $\mathcal{K}_u$ by the convex hull of $\mathcal{K}_r$ and the original $\mathcal{K}_u$. An alternative approach, which is perhaps more transparent, uses a family $\{\mathcal{K}_u(\alpha)\}$ of tolerance sets parameterized by confidence level α. For a given confidence level α and a given tolerance set $\mathcal{K}_r$, which may have been selected on the basis of the same confidence level, we then take $\mathcal{K}_u = \mathcal{K}_u(\beta)$ where β is defined by

$$\beta = \min(\alpha, \sup\{\gamma \in (0,1) : \mathcal{K}_r \in \mathcal{K}_u(\gamma)\}). \tag{11}$$

We will use the latter construction in the rest of the paper.

The analogs of Theorem 1 (invariance) and Theorem 2 (positive homogeneity) can easily be shown to hold for both estimation risk and misspecification risk separately. Proofs follow the lines of the proofs of the cited results.

IV. Application to portfolio risk management

One of the most important tasks of a risk management department is to compute the risk of the portfolio of the financial institution. The Bank for International Settlements (BIS) has suggested to set risk-based capital requirements which are closely related to the value-at-risk methodology. Here we show how our model risk measurement approach can be taken into account for a (simple) value-at-risk methodology.

Example 1 Value-at-Risk.
We have available a data set of returns $\left(h_1^T, ..., h_n^T\right)$ of the portfolio under consideration for a period of length nT years ($nT = 20$, $T = 1/252$ (one day)). An elementary VaR

model assumes that the data is a realization of a random sample $\left(H_1^T, ..., H_n^T\right)$ where $H_j^T \sim \mathcal{N}\left(\mu T, \sigma^2 T\right)$ for $j = 1, .., n$ where μ and σ^2 denote annualized mean and variance, respectively. We estimate μ by $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n h_i^T$ and σ by $\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(h_i^T - \hat{\mu}\right)}$. Let $\theta = (\mu, \sigma)$. The model class $\mathcal{M}$ with typical element

$$m(\theta) := \left\{ (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_\theta) : \mathbb{P}_\theta\left((-\infty, x]\right) = \Phi\left(\frac{\log x - \mu}{\sigma}\right) \right\}$$

is the class of the lognormal distributions.

Let $X_0 \in \mathbb{R}$ denote the (model independent) initial capital, and let $\Pi \in \mathcal{X}(\mathcal{M})$ denote the portfolio at time T. Formulas for various versions of value-at-risk are presented below; see the Appendix for derivations. To compute worst cases, we follow the approach of focusing directly on the given risk measurement method (VaR in this case) and the given product, as discussed in subsection B above.

First, assume that asset returns are normal and let $\hat{\theta}$ be an estimate of the parameter $\theta = (\mu, \sigma)$. We shall call $m(\hat{\theta})$ the *nominal parametric model*. The corresponding value-at-risk of portfolio Π at level p is given by

$$\mathrm{VaR}_{m(\hat{\theta})}\left(\Pi_{m(\hat{\theta})}\right) = -X_0 \exp\left(z_p \hat{\sigma} \sqrt{T} + \hat{\mu} T\right), \tag{12}$$

where z_p denotes the pth quantile of the (standard) normal distribution. Still assuming normality of asset returns, the parametric worst-case value-at-risk is the lower bound of the $(1 - \alpha)\%$ confidence interval around $\mathrm{VaR}_{m(\hat{\theta})}\left(\Pi_{m(\hat{\theta})}\right)$. This lower bound is given (based on an asymptotic approximation) by

$$\mathrm{VaR}_{m(\hat{\theta})}\left(\Pi_{m(\hat{\theta})}\right) - z_{\alpha/2}\sqrt{\Sigma_{\mathrm{VaR}}/n} \tag{13}$$

where
$$\Sigma_{\mathrm{VaR}} = T \, \mathrm{VaR}_{m(\hat{\theta})}\left(\Pi_{m(\hat{\theta})}\right)^2 \left[\sigma^2\left(1 + T/4\right) + z_p \sigma^3 \sqrt{T} + z_p^2 \sigma^2 / 2\right].$$

Nonparametric versions of VaR may be computed on the basis of the empirical distribution function F_n. We denote by m_n the model $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_n)$ where $\mathbb{P}_n$ is given by

$\mathbb{P}_n((-\infty, x]) = F_n(x)$. The nominal empirical value-at-risk is given by

$$\text{VaR}_{m_n}(\Pi_{m_n}) = -X_0 \exp\left(h^T_{n(\lfloor pn \rfloor + 1)}\right), \tag{14}$$

where $n(i)$ denotes the ith order statistic of $\left(h_1^T, ..., h_n^T\right)$ and $\lfloor a \rfloor$ is the largest integer that is less than or equal to a. Finally, the worst-case empirical VaR is the lower bound of the (nonparametric) confidence interval around $\text{VaR}_{m_n}(\Pi_{m_n})$, which may be computed as

$$\text{VaR}_{m_n}(\Pi_{m_n}) - z_{\beta/2}\sqrt{\frac{p(1-p)}{nf^2\left(F_n^{-1}(p)\right)}} \tag{15}$$

where β is as defined in (11) and where $f(x)$ can be estimated using, for instance, the Rosenblatt-Parzen kernel estimator.[5]

Example 2 Tail Conditional Expectation.

Below we use tail conditional expectation as well as value-at-risk, so we repeat the exercise of the previous example for TCE; again, see the Appendix for derivations. Assume the same setting as before. The nominal parametric TCE (at level p) under normality of the asset returns is given by

$$\text{TCE}_{m(\hat{\theta})}\left(\Pi_{m(\hat{\theta})}\right) = -\frac{1}{p}\exp\left(\hat{\mu}T + \tfrac{1}{2}\hat{\sigma}^2 T\right)\Phi\left(z_p - \hat{\sigma}\sqrt{T}\right). \tag{16}$$

The worst-case parametric TCE may be computed as

$$\text{TCE}_{m(\hat{\theta})}\left(\Pi_{m(\hat{\theta})}\right) - z_{\alpha/2}\sqrt{\Sigma_{TCE}/n} \tag{17}$$

---

[5]In our applications below we have approximately normal data and so we do bandwidth selection by taking $h = 1.06\sigma n^{-1/5}$, which is the optimal bandwidth in case of a normal $\mathcal{N}\left(\mu, \sigma^2\right)$ distribution.

where (writing $\phi = \Phi'$)

$$\Sigma_{\text{TCE}_{m(\theta)}} = T\Psi_{\text{TCE}}^2 \left[\sigma^2 + \left(T\sigma - \sqrt{T}\frac{\phi}{\Phi}\left(z_p - \sigma\sqrt{T} \right) \right) \sigma^3 \right.$$
$$\left. + \left(\sqrt{T}\sigma - \frac{\phi}{\Phi}\left(z_p - \sigma\sqrt{T} \right) \right)^2 \sigma^2/2 \right]. \quad (18)$$

The empirical TCE can be computed by

$$\text{TCE}_{m_n}\left(\Pi_{m_n} \right) = -\frac{1}{\lfloor np \rfloor + 1} \sum_{i=1}^n X_0 \exp\left(h_j^T \right) \mathbf{I}_{\{(-\infty, \text{VaR}_{m_n}(\Pi_{m_n})]\}}\left(h_j^T \right). \quad (19)$$

The empirical worst-case TCE can be computed as

$$\text{TCE}_{m_n}\left(\Pi_{m_n} \right) - z_{\alpha/2}\sqrt{\Sigma_{\text{TCE}_{mn}}/n} \quad (20)$$

where

$$\Sigma_{\text{TCE}_{mn}} = \frac{1}{p}\mathbb{E}_{\mathbb{P}_n}\left[Y^2 \mid Y \le \text{VaR}_{m_n}\left(\Pi_{m_n} \right) \right] - \left(\text{TCE}_{m_n}\left(\Pi_{m_n} \right) \right)^2$$
$$- \text{VaR}_{m_n}\left(\Pi_{m_n} \right)\left[\text{VaR}_{m_n}\left(\Pi_{m_n} \right) + \frac{2(1+p)}{p}\text{TCE}_{m_n}\left(\Pi_{m_n} \right) \right].$$

Figure 1 shows the normal density with variance equal to the sample variances (say, s^2) of the S&P 500 data and the British pound / US dollar (£/$) exchange rate data and compares this with a nonparametric density[6] estimate of the densities of the S&P 500 and the £/$ exchange rate. We see that the returns from the S&P 500 and an investment in British money market account exhibit more kurtosis than would be expected on the basis of normally distributed returns. This affects the value-at-risk and tail condition expectation computations as can be seen in Figures 2, 3, 4, and 5. Here we plot the 99% ($p = 0.01$) one-day value-at-risk and 95% ($p = 0.05$) tail conditional expectation for the S&P 500 and an investment in the British money market account. We choose the 99% level for VaR, since this is the quantile required by BIS (see Basle Committee

---

[6]In view of the approximate normality of the data, the bandwidth h has been set equal to $h = 1.06sn^{-1/5}$ which is the optimal bandwidth selection for normally distributed data.

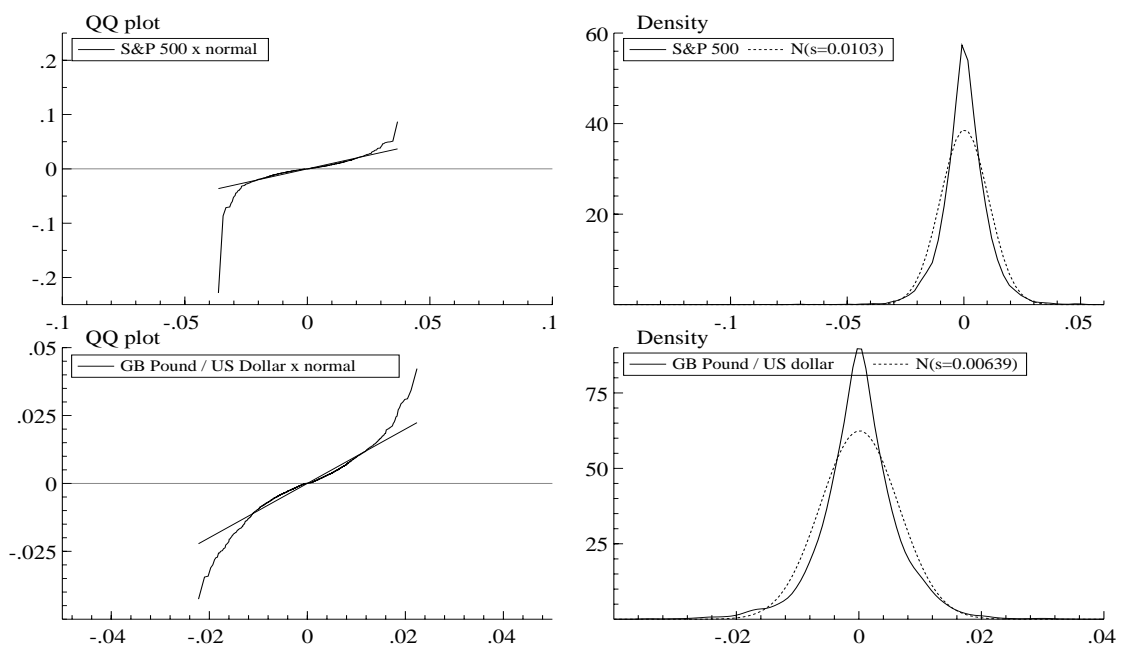**Figure 1.** QQ-plot and density comparison of the normal density with nonparametric density estimate (using a the Rosenblatt-Parzen kernel estimator with Gaussian kernel and bandwidth $h = 1.06sn^{-1/5}$) of the daily (total) returns of the S&P 500 and British pound / US dollar exchange rate. Data period 26-10-'81 – 26-10-'01 for the S&P 500 and 03-01-'86 – 26-10-'01 for the £ / $ exchange rate.

on Banking Supervision (1996)). For TCE we adopt a lower level, namely, the 95% level. The choice of level is a trade-off between having enough data points available for reliable estimation and putting enough weight on the extreme values. Since TCE automatically assigns more weight to extreme outcomes, the 5% level still provides a good representation of potential large losses while it increases the number of data points by about a factor of 5 compared to the 1% level. The volatility estimate is based on the past two years (500 days / data points) and is computed in the usual way. (For more advanced volatility estimation methods see, for example, Eberlein, Kallsen, and Kristen (2001).)

The relation between the worst-case risk measure and the risk measure based on the nominal model may be analyzed in terms of a multiplication factor. We define the multiplication factor for VaR or TCE as the ratio between the empirical worst-case VaR (TCE) based on a 95% confidence interval and the nominal parametric VaR (TCE). Plots of the multiplication factor in various cases are shown in the figures. We see that in case of 99% VaR multiplication factors of 2 for the S&P 500 and 1.6 for the £/$ exchange rate comfortably cover model risk at the 95% confidence level during the full sample period. In case of the 95% TCE we find that multiplication factors of 1.7 for the S&P 500 and 1.5 for the £/$ exchange rate are sufficient. For both the VaR and TCE a multiplication factor of three would correspond to a confidence level of 99.99%.

Ideally, the frequency of excessive losses (FOEL), that is the number of days at which the loss exceeds the predicted VaR, should be close to the VaR levels. In Tables I and II we present the results of a one-sided FOEL test at a level of 5. We denote by n the number of days in the backtesting period, by f the number of times the VaR level has been exceeded, and by $1-p$ the predicted level of VaR (97.5% or 99% in our case). The test statistic of the FOEL test is given by

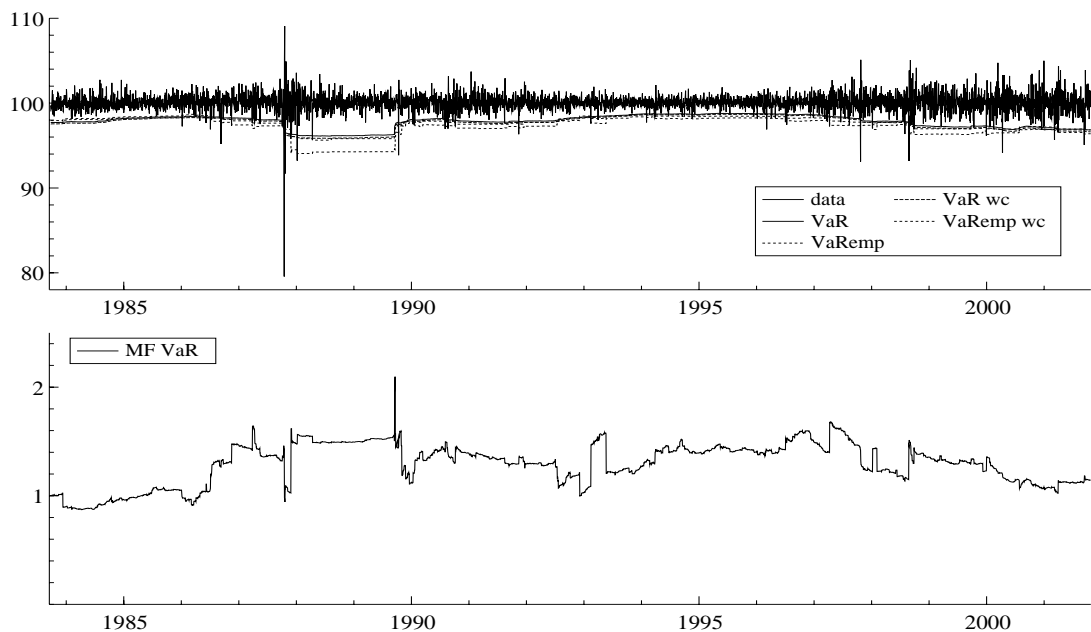$$T = \sqrt{n}\frac{f/n - p}{p(1-p)} \tag{21}$$

Figure 2. 99% VaR: The upper panel displays value-at-risk for the S&P 500 during the period 26-10-'81 – 26-10-'01. Shown are the nominal parametric VaR, the worst-case parametric VaR, the empirical VaR, and the empirical worst-case VaR. The lower panel shows the multiplication factor defined as the ratio between the empirical worst-case VaR (based on a 95% confidence interval) and the nominal parametric VaR.

Table I

FOEL test for nominal parametric VaR, worst-case parametric VaR, nominal empirical VaR and worst-case empirical VaR (for definitions, see main text). Daily data on S&P 500 (total return) index from 26-10-'81 to 26-10-'01.

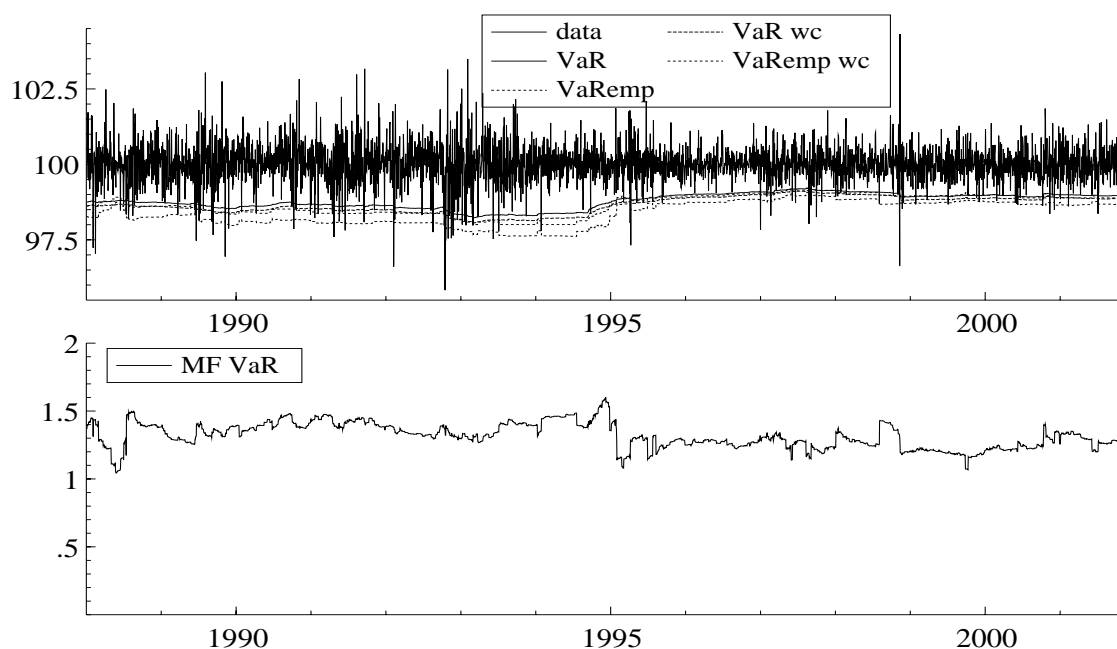|	VaR level	FOEL	1-sided 95% CI	T	p-value	VaR level hypothesis rejected
nom. par.	97.5%	3.2%	(2.8%; −)	2.7	0.00	yes
par. wc	97.5%	2.6%	(2.2%; −)	0.5	0.32	no
nom. emp.	97.5%	3.0%	(2.6%; −)	2.0	0.02	yes
emp. wc	97.5%	2.1%	(1.7%; −)	−1.9	0.97	no
nom. par.	99%	2.1%	(1.7%; −)	5.2	0.00	yes
par. wc	99%	1.7%	(1.4%; −)	3.8	0.00	yes
nom. emp.	99%	1.6%	(1.2%; −)	3.1	0.00	yes
emp. wc	99%	1.0%	(0.7%; −)	−0.3	0.63	no

23

Figure 3. 99% VaR: The upper panel displays the value-at-risk for the British pound / US dollar exchange rate during the period 03-01-'86 – 26-10-'01. Shown are the nominal parametric VaR, the worst-case parametric VaR, the empirical VaR, and the empirical worst-case VaR. The lower panel shows the multiplication factor defined as the ratio between the empirical worst-case VaR (based on a 95% confidence interval) and the nominal parametric VaR.

Table II

FOEL test for nominal parametric VaR, worst-case parametric VaR, nominal empirical VaR and worst-case empirical VaR (for definitions, see main text). Daily data on British pound / US dollar from 03-01-'86 to 26-10-'01.

	VaR level	FOEL	1-sided 95% CI	T	p-value	VaR level hypothesis rejected
nom. par.	97.5%	3.3%	(2.6%; −)	2.6	0.00	yes
par. wc	97.5%	2.8%	(1.9%; −)	0.0	0.49	no
nom. emp.	97.5%	2.6%	(2.0%; −)	0.3	0.40	no
emp. wc	97.5%	1.6%	(0.9%; −)	−6.2	1.00	no
nom. par.	99%	1.9%	(1.4%; −)	4.0	0.00	yes
par. wc	99%	1.4%	(0.8%; −)	1.4	0.07	no
nom. emp.	99%	1.2%	(0.7%; −)	0.9	0.19	no
emp. wc	99%	0.6%	(0.2%; −)	−4.6	1.00	no
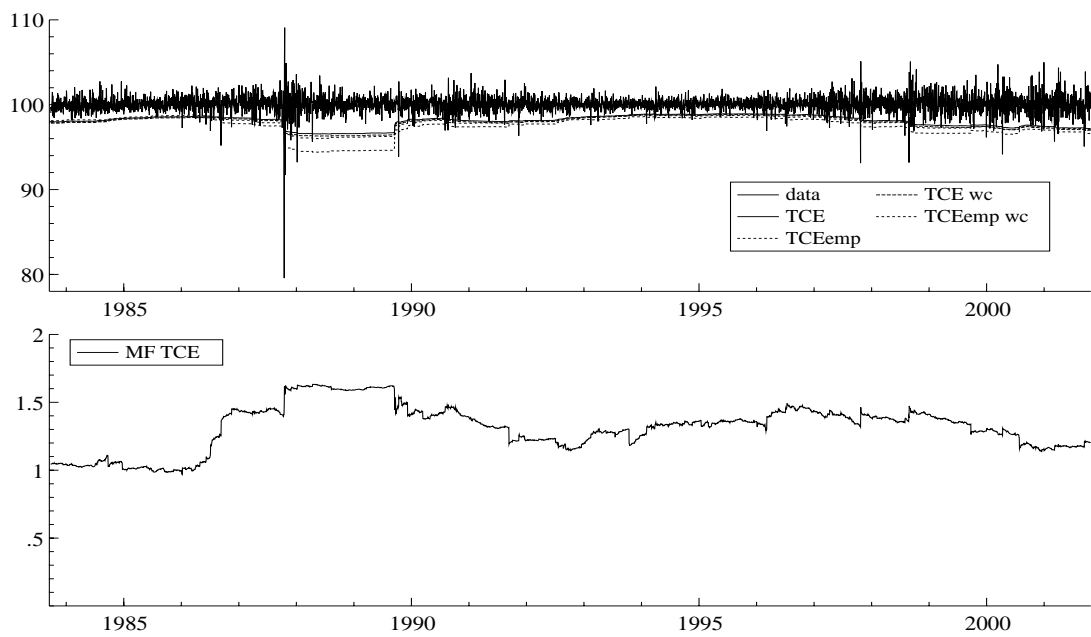
Figure 4. 95% TCE: The upper panel displays the tail conditional expectation for the S&P 500 during the period 26-10-'81 – 26-10-'01. Shown are the nominal parametric TCE, the worst-case parametric TCE, the empirical TCE, and the empirical worst-case TCE. The lower panel shows the multiplication factor defined as the ratio between the empirical worst-case TCE (based on a 95% confidence interval) and the nominal parametric TCE.
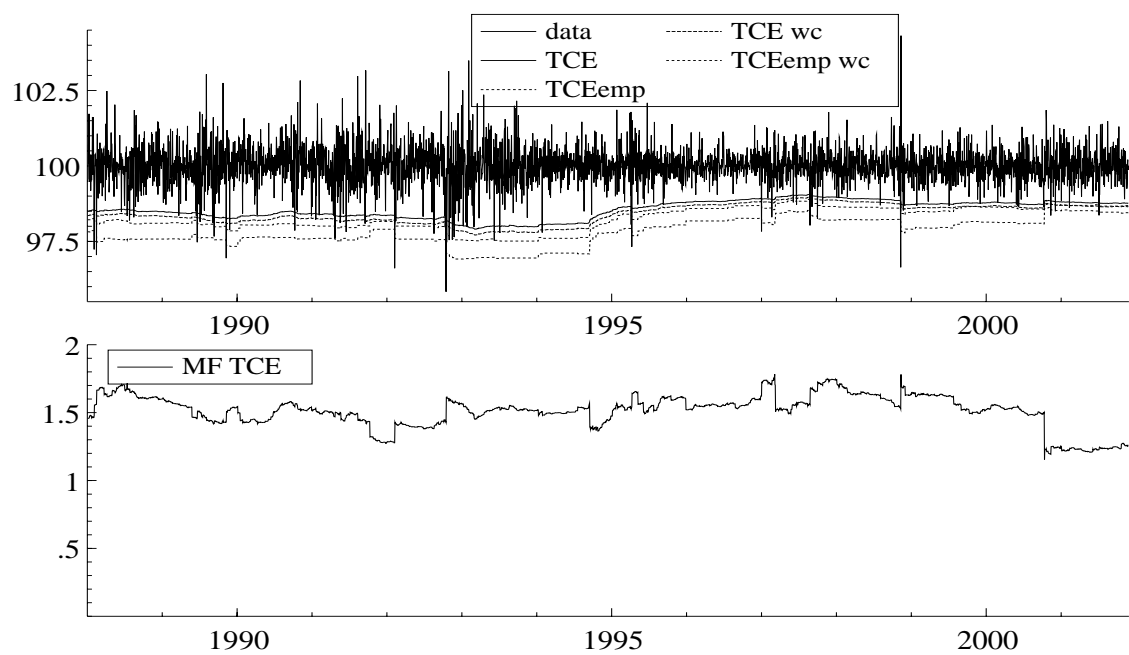
Figure 5. 95% TCE: The upper panel displays the tail conditional expectation for the British pound / US dollar exchange rate during the period 03-01-'86 – 26-10-'01. Shown are the nominal parametric TCE, the worst-case parametric TCE, the empirical TCE, and the empirical worst-case TCE. The lower panel shows the multiplication factor defined as the ratio between the empirical worst-case TCE (based on a 95% confidence interval) and the nominal parametric TCE.

26

The results indicate that the VaR model based on independent normally distributed returns is strongly rejected both in case of the S&P 500 data and in case of the £/$ data. For both the S&P 500 and the £/$ exchange rate, taking estimation risk into account seems sufficient in case of the 97.5% level since we cannot reject the worst-case normal models for this quantile. In case of the 99% level, however, taking estimation risk into account does not prevent the VaR limit from being exceeded too often. If we take misspecification risk into account by looking at the empirical worst-case model, the number of times the VaR limit is crossed does not exceed the level predicted by the model in a statistically significant way.

V. Application to derivative securities

In this section we want to determine a model reserve for the models that derivative traders use. We extend the analysis of Green and Figlewski (1999) by explicitly basing our model reserve on a risk measure. Before we can determine this model reserve, we should first define the risk associated with a derivative security. Theoretically, derivative assets can be exactly replicated (in case of a complete market) by a (dynamic) position in the underlying asset and some numeraire asset. In practice, these replicating strategies are not feasible due to transaction costs and the inability to trade continuously. Therefore, financial institutions rely on hedge strategies in discrete time. By hedging in discrete time, however, the position consisting of the derivative and the hedging portfolio is no longer risk free and subjected to market risk. We use this market risk to get to a definition of the risk associated with a derivative. The model reserve can then be based on the misspecification of the market risk associated with the hedged derivative.

A. Derivative risks

For a portfolio of basis assets the value of the portfolio at the relevant time point is a natural definition of the market risk of that portfolio. In principle, the same can be done for derivatives. However, this would ignore the fact that financial institutions have the possibility (and make heavily use of this possibility) to hedge their derivative

portfolio. Indeed, Green and Figlewski (1999) found that derivative risks can be reduced considerably by delta hedging. Therefore, it is wise to take the hedge strategy of the financial institution into account when calculating the market risk of a derivative. We suggest to define the market risk of a derivative as the market risk of the cost of hedging of this derivative. The cost of hedging $C(X; \gamma)$ of claim X using trading strategy γ is given by [7]

$$C(X; \gamma) \equiv \frac{X}{N}(T) - \sum_{i=1}^{n} \gamma(t_{i-1}) \cdot \Delta\left(\frac{S}{N}\right)(t_i) - \sum_{i=1}^{n} \gamma(t_{i-1}) \cdot \frac{\delta \odot S}{N}(t_i)(t_i - t_{i-1}), \quad (22)$$

where S denotes the underlying asset(s), N denotes a numeraire asset, δ denotes the dividend process, and $\Delta\left(\frac{S}{N}\right)(t_i) \equiv \left(\frac{S}{N}\right)(t_i) - \left(\frac{S}{N}\right)(t_{i-1})$.

For a formal definition we should introduce a filtration such that the underlying process and the trading strategy are well defined w.r.t. this filtration. However, since we are only interested in (discounted) final payoffs of the derivative we do not need to extend the framework in Section II.B. Using a trading strategy γ we can introduce a new risk $Y = C(X; \gamma)$ defined on m. If we would also be interested in payoffs during the life of the option, we would need to extend the framework in Section II.B by using processes instead of random variables.

In principle, the trading strategy γ used does not have to be related to the pricing model. However, for ease of exposition we restrict our attention to the delta hedge which is the most commonly used hedge technique in practice. Of course, a basis asset can also be seen as a contingent claim.

B. Estimation risk and misspecification risk

In most financial institutions a trading desk (or its research department) should provide risk management with a pricing model and a hedge (trading) strategy for the derivative they would like to trade. The task of risk management is then to estimate the market risk associated with this pricing model and hedge strategy. Consider the following example.

---

[7]The symbol $\odot$ denotes the Hadamard product, that is, $x \odot y = (x_1 y_1, ..., x_n y_n)$ (see, for example, Magnus and Neudecker (1999)).

The trading desk would like to trade plain vanilla (call, put) options on equity and exchange rates. To do so they suggest to price the options using the Black-Scholes model with implicit volatilities (which basically means taking the market prices as given). The hedge strategy suggested is a daily delta hedge where the delta is computed using the BS model with this implicit volatility.

To get an estimate of the market risk of the derivative using definition (22) the financial institution needs to get an estimate of the risk profile (cdf) of the cost of hedging. If the model suggested is correct this estimate can be obtained by a Monte Carlo simulation of the dynamics of the underlying and computing the cost of hedging using the pricing and hedge formulas. By increasing the number of simulations the desired accuracy level can be reached. However, this would lead to the market risk of the derivative in a BS model. To take possible model misspecification into account we need to see how the model performs on empirical data.

Since we obviously cannot perform a Monte Carlo simulation from the actual dynamics we rely on historical simulation, following, among others, Galai (1977), Merton, Scholes, and Gladstein (1978), Merton, Scholes, and Gladstein (1982), and Green and Figlewski (1999). To perform the historical simulation we have available a data set of the underlying $\{s_0, ..., s_n\}$ and from these we get the (daily) logreturns $\{h_1, ..., h_n\}$. To compute the implicit volatilities, we should have data on the option prices. The data should cover a large range of different moneyness levels, since over time the options written move in and out of the money. Since we do not have data on option prices available we approximate the option prices by computing them using the BS formula with a one year historical volatility estimator which is also used to compute the hedge ratios (BS deltas).[8] This approach has the advantage that we can write a new option every day. We start by writing an option f at time $t = 0$ and continue this until $t = n - k$ where $k = T * no$, and T denotes the maturity of the option in years and no is the number of days in a year. In addition to the option prices $\{f(0), ..., f(n-k)\}$

---

[8]Though implicit volatility is by definition superior to historical volatility for pricing purposes, its superiority for hedging purposes is unclear. In a review article Figlewski (1997) shows that typically implicit volatilities contain only marginally more information about future volatility than historical volatility estimators.

we compute the actual cost of hedging the option using the specified hedge strategy to get $\{C(0), ..., C(n-k)\}$ (see eq. (22)). Using this data we compute an estimator for the cdf of the (discounted) final profit and loss ($P\&L$) account $P\&L \equiv f - C$. The cdf (or pdf) of $P\&L$ can be seen as the return distribution of pricing derivatives with the specified pricing model (in our case the BS model with historical volatility) and the specified hedge strategy (daily hedge based on BS deltas using historical volatility). On this distribution we can then compute the relevant risk measures described in Section II.C.

In doing so we need to take into account the fact that the data is subject to the overlapping samples problem. We handle this problem by using the method of Newey and West (1987). We can compare the computed empirical risk measure to the risk measure obtained by the Monte Carlo simulation from the model to see the impact of model risk. However, we cannot judge whether the difference is due to model misspecification or estimation risk since the precisions of the estimates of the risk measure obtained by the Monte Carlo simulation and the empirical risk measure are not (necessarily) equal. To analyze the part of the model risk due to model misspecification and estimation risk, we perform an auxiliary computation. We generate an auxiliary data set of returns $\{h_1^*, ..., h_n^*\}$ and prices $\{s_0^*, ..., s_n^*\}$ according to the model assumptions of equal length to the original sample. Since in the BS model the returns are normally distributed, we generate h_i^* from $\mathcal{N}\left(\overline{h}_n, \frac{1}{n}\sum_{i=1}^{n}\left(h_i - \overline{h}_n\right)^2\right)$.[9] On this auxiliary data we perform the same analysis as on $\{s_0, ..., s_n\}$ and $\{h_1, ..., h_n\}$. The difference between the empirical data and the model data lies therefore in the fact that we removed the dependence structure from $\{h_1, ..., h_n\}$ and that all moments in the auxiliary data set are determined by the first two.

C. Set-up of experiment

We investigate two major markets on which financial options are actively traded: The Standard and Poor's 500 (SPX) for equity options and the British pound / U.S. dollar

---

[9]To reduce sampling error we corrected $\{h_1^*, ..., h_n^*\}$ such that $\overline{h}_i^* = \overline{h}_n$ and $\frac{1}{n}\sum_{i=1}^{n}\left(h_i^* - \overline{h}_n\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(h_i - \overline{h}_n\right)^2$.

(£/$) foreign exchange (FX) rate. Sample periods vary depending on data availability. The data was obtained from Datastream (definitions and sources of the data can be found in Appendix B). The models that we investigate all come from the Black-Scholes (BS) framework. Though its deficiencies have been demonstrated numerous times, it remains the most widely used option pricing model.

The original BS model (see Black and Scholes (1973)) was designed for European options on non-dividend paying stock. Since the SPX consists of dividend paying stocks, we use the adjusted BS model by Merton (1973) which allows for a continuous proportional dividend yield δ. Since future dividends are unknown, this represents another source of risk in trading derivatives, namely dividend risk. We compute the option prices using the realized dividend yield.[10] For FX options we use the Garman-Kohlhagen model (see Garman and Kohlhagen (1983)) which adjusts the original BS model for options on foreign currencies. For the pricing formulas we refer to the original papers, Green and Figlewski (1999), or Hull (2000).

To compare risks of derivatives with different characteristics (call/put flag, moneyness, and time to maturity) we always write enough contracts to generate a premium (initial value) of $100. The results can then be interpreted as a dollar return on an investment of $100 in the specified contract or percentage returns per dollar of option premium. To limit derivatives risk one can essentially distinguish three strategies. The first strategy is to diversify using derivatives with different characteristics and other risky assets. The second approach is using cash flow matching which consists of creating offsetting positions with different counterparties such that the derivatives contract is replicated. Though cash flow matching is the most precise method of hedging and, furthermore, model independent, it is rarely possible for a financial institution to construct a cash flow matching hedge. In general, the public wants to be long in options which brings about the short options position of the financial industry and, thereby, making it impossible for the financial institution to match all of its cash flows. Finally,

---

[10]We also did the analysis using the expected dividend yields from Datastream (for the definitions and computations, see Appendix C) which resulted only in minor changes. Furthermore, we performed the analysis on the S&P500 total return index as if the S&P 500 was a non-dividend paying index. This also resulted in minor changes.

the financial institution can hedge using delta hedging[11] to hedge the derivatives risk. Since cash flow matching is impractical and Green and Figlewski (1999) showed that delta hedging is far superior to hedging by diversification, we restrict ourselves to delta hedging which is also the most often used hedge strategy in the financial industry.

ITM options end in the money much more than OTM options. This could result in differences in the effectiveness of the hedge strategy. To investigate the influence of moneyness we compare the results for ITM, ATM, and OTM options where we take OTM options with a strike of 0.4 standard deviations below or above (ITM or OTM) the mean.[12] This more or less covers the range of options seen in the market.

D. Results

In this section we discuss the results from the experiment described above. Figure 6 shows the nonparametric estimates of the densities of the $P\&L$ for a one year at-the-money[13] (ATM) call option, put option, and straddle on the S&P 500. We see that the differences are small. Since traders prefer to trade straddles (volatility trading) rather than single calls or puts (taking a market view) from now on. We see that the density estimates under the model assumptions are more or less symmetric, while the empirical density estimates are skewed to the left. Therefore, it happens much more often than expected that the actual cost of hedging exceeds the option premium by a substantial amount. To measure the risk it is therefore wise to consider risk measures taking the tail into account such as the VaR and TCE.

In Table III we present the mean return, 99% VaR, 95% VaR, 95% TCE, and 90% TCE of a portfolio of straddles with an initial value of $100 on the real data. We see that on average this premium of $100 is about the amount needed to hedge the derivative. Except for the 2 year £/$ exchange rate none of the mean returns differs statistically significantly from zero. Due to the overlapping samples that arise if we write

---

[11]Traders also often hedge other greeks (gamma, vega, etc.). As argued by Green and Figlewski (1999) this requires, however, other options which need to be bought from other financial institutions. The overall financial industry is therefore restricted to delta hedging.

[12]To put it more precisely, $k^{\pm}(T) = F(t)\exp\left(\pm 0.4\sigma\sqrt{T-t}\right)$ with $k^{-}(T)$ the OTM strike with maturity date T and $k^{+}(T)$ the ITM strike with maturity date T.

[13]Moneyness is defined as $m = \log(F/k)$ where F denotes the futures price and k the strike price.
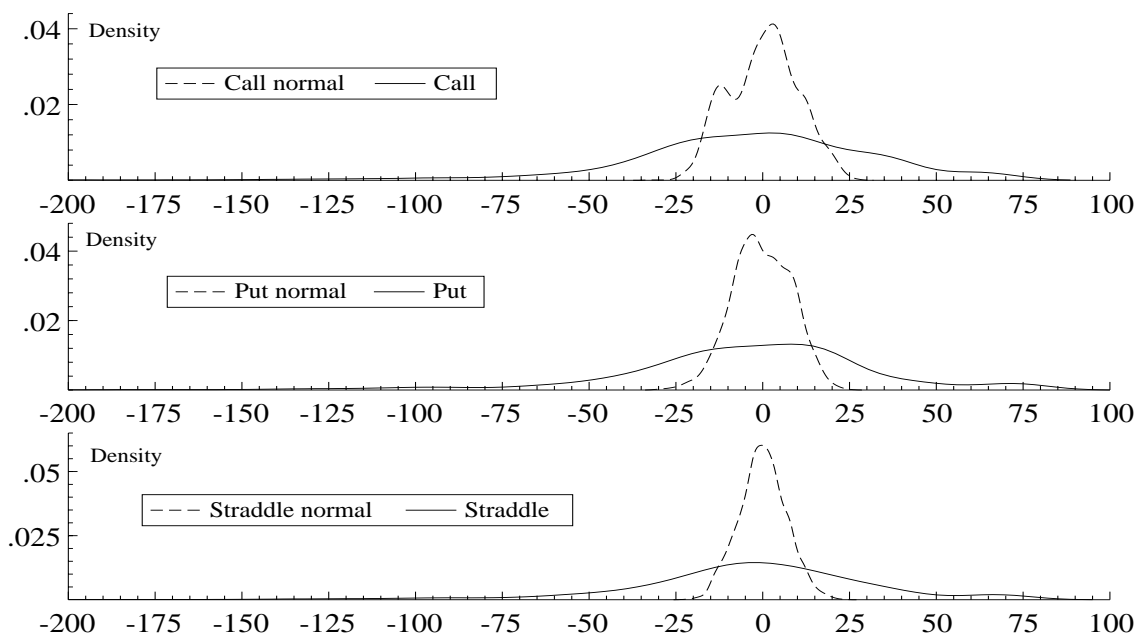
Figure 6. Plot of nonparametric estimates of the density of the cost of hedging of a 1 year ATM option on the S&P 500. The cost of hedging for the Black-Scholes model (normal), a worst-case Black-Scholes model, and the empirical cost of hedging is plotted. In the upper panel a call is plotted, in the middle panel a put, and in the lower panel a straddle.
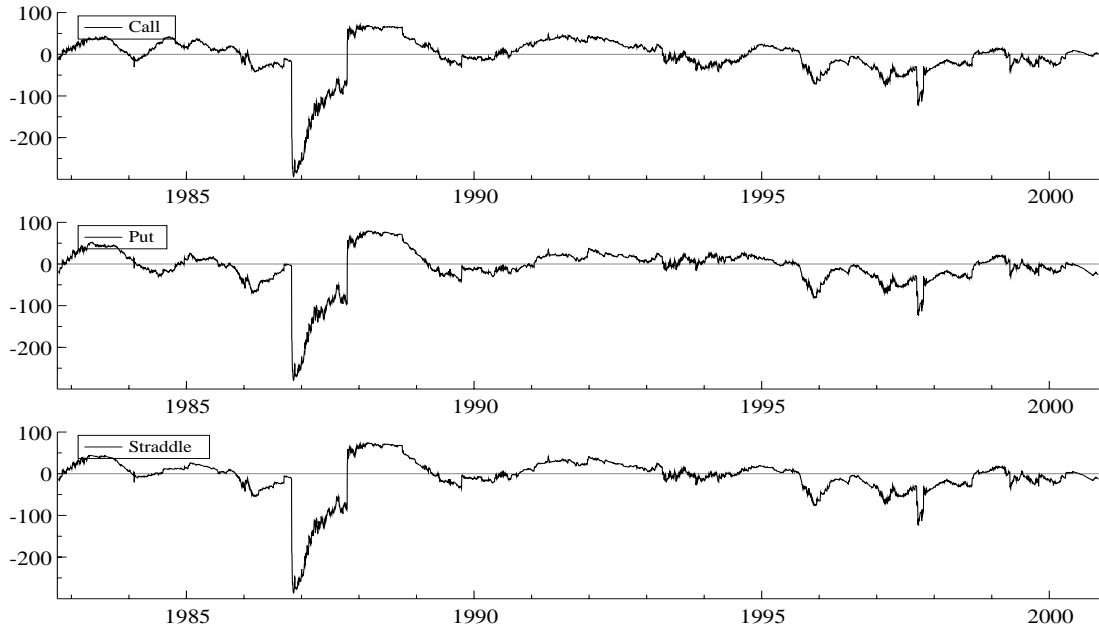
Figure 7. Time series of the cost of hedging over time for call, put, and straddle in the upper, middle, and lower panel, respectively.

a new option contract every day, we need to correct the standard errors. For example, in Figure 7 we see that the costs of hedging are highly positively correlated over time. This correlation can be taken into account by using the Newey and West (1987) standard errors with a period length equal to the time to maturity minus one day.[14] The difference between regular standard errors and Newey and West (1987) standard errors is found to be quite substantial. Due to the dependence of our model risk measure on a worst case market risk measure, this difference is important for the calculation of the model risk and should be taken into account.[15] Consider, for example, the 5% VaR of the $P\&L$ for the straddle in the upper panel with a maturity of 1 year. The estimated total risk is given by $86.4 + 1.96 * 47.2 = 178.9$ whereas it would have been $86.4 + 1.96 * 4.2 = 94.6$ with the use of regular standard errors which is an underestimation of the total risk of

---

[14] After the time to maturity minus one day the theoretical autocorrelation is zero. However, often it is recommended to take some more lags into account since the longer lags are quite substantially downweighted in the Newey-West standard errors. We included some extra lags and found more or less the same results and therefore choose to restrict to taking time to maturity minus one day.

[15] This was not done in Green and Figlewski (1999) and leads to underestimation of the model risk.

34

53%. In the second panel of Table III we use the data from 01-01-'88 to 26-10-'01 to investigate the impact of the crash of 1987 on our results. We see that the VaRs and TCEs decrease dramatically for the 1 and 2 year maturity. The reason that the results for the one and three months options are more or less the same (the standard errors are obviously bigger due to fewer data) is that less observations were influenced by the crash due to their shorter time to maturity. Comparing the results including and excluding the crash, we notice that, if we include the crash, the option positions seem riskier the longer the time to maturity. Excluding the crash produces the opposite result, that is, the longer maturity positions seem less risky. In case of the £/$ exchange rate there does not seem to be a clear dependence on time to maturity.

Tables IV and V present a decomposition of the total risk consisting of the market risk computed according to the model and the model risk which itself can be decomposed into estimation risk and misspecification risk. For example, if we take the 1 year ATM straddle on the S&P 500 (using the '88-'01 sample) we find that of the total risk according to the 5% TCE of $89.6 consists for 49% ($44.1) of market risk due to discrete hedging. This risk could be estimated by using the model. The estimation risk is 16% ($13.9) which could also be found by using the model and calculating the appropriate standard errors. The misspecification risk is substantial and amounts to 35% ($31.6). It is clear that the model risk due to misspecification risk prevails over the estimation risk. Furthermore, we see that the model estimates the market risk induced by discrete hedging at a much too low level.

In Tables VI and VII we study the effect of the moneyness on the performance of the model. Since straddles consist of a call and a put with the same strike, we give separate results for calls and puts. For the definition of the moneyness, we took the moneyness of the call option of the straddle. We first see for OTM options that call and put options, separately are much riskier than a straddle consisting of an OTM call and an ITM put. Therefore, the risk of writing an OTM call (or put, see results for ITM straddle) can be substantially decreased by adding a put (or call) to the portfolio. For the ATM and ITM options the results are more or less the same for calls, puts, and straddles. Comparing

35

Table III: **Return, VaR, and TCE of the** $P\&L$ **of writing options**

The upper panel of the table reports the performance of the delta hedge strategy for ATM straddles using a historical volatility estimator based on 1 year data. Between brackets both the Newey-West standard errors correcting for overlapping samples (on the left) and regular standard errors (on the right) are reported.

Underlying	Maturity	Mean Return	1% VaR	5% VaR	5% TCE	10% TCE	Percentage calls In-the-money
SPX	1 month	1.69 (1.96; 0.51)	−108.1 (20.4; 7.4)	−58.2 (5.2; 1.8)	−94.9 (11.6; 3.9)	−70.9 (7.4; 2.4)	59%
'81-'01	3 months	−0.11 (3.20; 0.47)	−103.6 (11.6; 3.2)	−57.5 (6.3; 1.3)	−83.9 (10.4; 2.4)	−66.3 (7.8; 1.6)	66%
	1 year	−7.93 (8.32; 0.68)	−245.6 (44.4; 6.9)	−86.4 (47.2; 4.2)	−156.8 (59.7; 6.2)	−109.2 (43.4; 4.0)	77%
	2 years	8.21 (11.49; 0.80)	−193.3 (50.9; 8.3)	−79.8 (48.7; 3.7)	−150.6 (62.7; 5.5)	−106.8 (44.4; 3.5)	87%
SPX	1 month	−0.78 (2.28; 0.61)	−119.6 (37.3; 12.7)	−58.5 (6.1; 2.1)	−97.6 (16.1; 4.9)	−72.6 (9.9; 3.0)	59%
'88-'01	3 months	−3.36 (3.45; 0.52)	−106.5 (11.4; 3.3)	−60.6 (9.0; 1.9)	−89.4 (13.6; 2.9)	−70.4 (10.4; 2.0)	67%
	1 year	−6.93 (5.65; 0.44)	−77.3 (13.5; 3.1)	−51.8 (10.0; 1.3)	−68.8 (10.6; 1.9)	−56.6 (11.1; 1.4)	81%
	2 years	−10.74 (10.47; 0.61)	−58.7 (3.7; 0.8)	−47.2 (11.4; 1.2)	−55.1 (6.4; 0.8)	−47.6 (9.6; 0.9)	89%
£ / $	1 month	−0.03 (2.60; 0.69)	−132.7 (45.5; 13.4)	−68.1 (6.5; 2.2)	−118.4 (22.4; 6.6)	−86.7 (13.1; 3.8)	54%
'86-'01	3 months	−1.86 (4.05; 0.63)	−132.5 (34.0; 6.8)	−74.2 (11.7; 2.3)	−108.2 (18.7; 3.4)	−84.7 (13.6; 2.4)	53%
	1 year	1.73 (5.01; 0.44)	−66.7 (9.2; 2.4)	−36.1 (7.6; 1.2)	−54.4 (9.4; 1.7)	−42.0 (7.5; 1.2)	54%
	2 years	38.73 (12.97; 1.27)	−136.1 (11.6; 2.9)	−91.4 (29.0; 3.9)	−119.4 (18.3; 2.6)	−92.5 (23.0; 2.9)	64%

Table IV
Decomposition of Model Risk for 1% VaR

In this table the total risk measured by the 1% VaR of a straddle is split into three components: The market risk due to discrete hedging (according to the model), the estimation risk involved in determining this market risk, and the misspecification risk. For every maturity the first row reports the cumulative absolute total risk, while in the second row the percentages of the total risk are shown.

1% VaR

Underlying	Maturity	market risk	est. risk	misspec. risk	Percentage calls In-the-money
SPX	1 month	−56.8	−66.1	−148.1	59%
		38%	6%	56%	
'81-'01	3 months	−33.3	−37.4	−126.3	66%
		27%	3%	70%	
	1 year	−15.2	−16.4	−332.6	77%
		5%	0%	95%	
	2 years	−17.5	−25.1	−293.0	87%
		8%	3%	89%	
SPX	1 month	−88.6	−106.0	−192.8	59%
		46%	9%	45%	
'88-'01	3 months	−85.1	−129.1	−129.1	67%
		66%	34%	0%	
	1 year	−50.9	−64.8	−103.8	81%
		49%	13%	38%	
	2 years	−41.6	−47.2	−148.1	89%
		63%	8%	29%	
£/$	1 month	−58.9	−65.3	−221.8	54%
		26%	3%	71%	
'86-'01	3 months	−58.3	−65.4	−199.1	53%
		29%	4%	67%	
	1 year	−26.1	−34.3	−84.8	54%
		31%	10%	59%	
	2 years	−90.1	−98.8	−158.9	64%
		57%	5%	38%	

Table V
Decomposition of Model Risk for 5% TCE

In this table the total risk measured by the 5% TCE of a straddle is split into three components: The market risk due to discrete hedging (according to the model), the estimation risk involved in determining this market risk, and the misspecification risk. For every maturity the first row reports the cumulative absolute total risk, while in the second row the percentages of the total risk are shown.

5% TCE

Underlying	Maturity	market risk	est. risk	misspec. risk	Percentage calls In-the-money
SPX	1 month	−47.4	−53.0	−117.6	59%
		40%	5%	55%	
'81-'01	3 months	−29.9	−33.9	−104.4	66%
		28%	4%	68%	
	1 year	−14.1	−16.2	−273.8	77%
		5%	1%	94%	
	2 years	−19.7	−24.6	−273.5	87%
		7%	2%	91%	
SPX	1 month	−73.9	−87.5	−129.3	59%
		57%	11%	32%	
'88-'01	3 months	−67.2	−88.8	−116.1	67%
		58%	19%	23%	
	1 year	−44.1	−58.0	−89.6	81%
		49%	16%	35%	
	2 years	−37.9	−45.6	−67.7	89%
		56%	11%	33%	
£/$	1 month	−52.9	−57.9	−162.2	54%
		33%	3%	64%	
'86-'01	3 months	−52.4	−60.0	−144.7	53%
		36%	5%	59%	
	1 year	−21.5	−33.2	−72.9	54%
		30%	16%	54%	
	2 years	−85.5	−95.2	−155.3	64%
		55%	6%	39%	

Table VI: **1% VaR for different types and moneyness**

This table reports the market risk (according to the model), estimation risk, and misspecification risk for OTM, ATM, ITM calls, puts, and straddles measured by the 1% VaR. The first row gives the percentage of options that ended in the money. In the second row the cumulative absolute total risk in dollar returns on a $100 investment in the option is shown. The third row reports the percentages of the total risk due to 1) market risk, 2) estimation risk, and 3) misspecification risk.

| | | 1% VaR | | | | | | | | |
| | | OTM | | | ATM | | | ITM | | |
Underlying	Type	market risk	est. risk	misspec. risk	market risk	est. risk	misspec. risk	market risk	est. risk	misspec. risk
SPX	call		59%			77%			85%	
		−30.5	−33.3	−979.0	−20.8	−24.3	−338.7	−14.7	−17.7	−157.9
		3%	0%	97%	6%	1%	93%	9%	2%	89%
	put		14%			23%			41%	
'81-'01		−29.5	−35.2	−450.2	−20.3	−25.1	−329.8	−17.0	−22.6	−333.4
		7%	1%	92%	6%	1%	93%	5%	2%	93%
	straddle		59%			77%			85%	
		−12.7	−14.8	−494.6	−15.2	−16.4	−332.6	−17.5	−25.1	−227.9
		3%	0%	97%	5%	0%	95%	8%	3%	89%
SPX	call		59%			81%			88%	
		−102.8	−120.9	−278.8	−57.1	−71.9	−95.5	−29.7	−33.8	−37.7
		37%	6%	57%	60%	15%	25%	79%	11%	10%
	put		12%			19%			41%	
'88-'01		−71.4	−78.3	−156.6	−44.6	−49.5	−116.6	−33.5	−37.5	−51.3
		46%	4%	50%	38%	4%	58%	65%	8%	27%
	straddle		59%			81%			88%	
		−50.9	−57.5	−84.5	−50.9	−64.8	−103.8	−39.8	−45.4	−55.6
		60%	8%	32%	49%	13%	38%	72%	10%	18%
£/$	call		35%			54%			75%	
		−57.1	−65.0	−289.7	−43.1	−59.4	−81.0	−35.1	−54.1	−54.2
		20%	3%	77%	53%	20%	27%	65%	35%	0%
	put		25%			−46%			67%	
'86-'01		−60.3	−81.9	−217.6	−37.4	−45.4	−97.1	−25.3	−32.2	−49.6
		28%	10%	62%	39%	8%	53%	51%	15%	34%
	straddle		33%			54%			75%	
		−20.6	−25.8	−92.2	−26.1	−34.3	−84.8	−21.6	−28.6	−64.4
		22%	6%	72%	31%	10%	59%	34%	10%	56%

Table VII: **5% TCE for different types and moneyness**

This table reports the market risk (according to the model), estimation risk, and misspecification risk for OTM, ATM, ITM calls, puts, and straddles measured by the 5% TCE. The first row gives the percentage of options that ended in the money. In the second row the cumulative absolute total risk in dollar returns on a $100 investment in the option is shown. The third row reports the percentages of the total risk due to 1) market risk, 2) estimation risk, and 3) misspecification risk.

		OTM			5% TCE ATM			ITM		
Underlying	Type	market risk	est. risk	misspec. risk	market risk	est. risk	misspec. risk	market risk	est. risk	misspec. risk
SPX	call		59%			66%			85%	
		−28.0	−31.0	−450.0	−18.6	−21.6	−289.4	−13.1	−16.2	−173.2
		6%	1%	93%	6%	1%	93%	8%	2%	90%
'81-'01	put		14%			23%			41%	
		−26.1	−31.4	−448.2	−17.7	−20.9	−258.2	−14.4	−18.4	−133.2
		6%	1%	93%	7%	1%	92%	11%	3%	86%
	straddle		59%			66%			85%	
		−11.5	−13.2	−211.7	−13.8	−16.2	−273.8	−14.0	−17.8	−243.4
		5%	1%	94%	5%	1%	94%	6%	2%	92%
SPX	call		59%			81%			88%	
		−91.7	−112.0	−252.4	−49.5	−68.2	−90.3	−27.0	−35.8	−38.0
		36%	8%	56%	54%	21%	24%	71%	23%	6%
'88-'01	put		12%			19%			41%	
		−66.6	−76.6	−138.0	−42.2	−48.8	−89.0	−30.9	−36.3	−46.4
		48%	7%	45%	47%	7%	46%	66%	12%	22%
	straddle		59%			81%			88%	
		−46.4	−54.9	−76.3	−44.1	−58.0	−89.6	−35.6	−47.1	−55.7
		61%	11%	28%	49%	16%	35%	64%	21%	15%
£/$	call		35%			54%			75%	
		−51.2	−61.8	−207.2	−37.6	−48.5	−72.7	−28.7	−39.8	−44.9
		25%	5%	70%	52%	15%	33%	64%	25%	11%
'86-'01	put		25%			−46%			67%	
		−48.8	−72.0	−160.1	−31.9	−44.1	−74.9	−21.3	−26.4	−44.1
		31%	14%	55%	43%	16%	41%	48%	12%	40%
	straddle		33%			54%			75%	
		−17.9	−23.8	−64.8	−21.5	−33.2	−72.9	−18.6	−23.4	−54.7
		28%	9%	63%	30%	16%	54%	34%	9%	57%

the riskiness of the options with different levels of moneyness, we notice that the OTM options are riskier than the ITM options. The risk of the ATM options lies in between. Another interesting feature is that for OTM options the misspecification risk is more important than for ITM options. This indicates that ITM options are more easy to hedge. Once an option is clearly ITM, the delta is close to one and it does not make much difference whether the underlying data is more skewed or fat-tailed.

VI. Conclusions

In this paper we have presented a framework to set capital requirements for trading activities in a market based on the extent to which this market can be reliably modeled. Our framework extends the (market) risk framework set out by Artzner, Delbaen, Eber, and Heath (1999) and Delbaen (2000) by considering risk measurement methods for a class of models instead of a risk measure for one particular model. This allows for a quantification of model risk on top of market risk measurement.

The general framework presented is elaborated in such a manner that it fits well in the capital adequacy framework set out by the Basle Committee and that of many internal risk management divisions. The use of risk measurement methods extends the currently used value-at-risk and the recently proposed coherent risk measures in a natural way.

We decompose the total model risk into a component due to estimation error and a component due to misspecification. This is established using a tolerance set restricted to a model class in order to quantify estimation risk and an unrestricted tolerance set to quantify misspecification risk. This allows a division of capital requirements currently used (for example, the multiplication factor of the BIS) in market risk, model risk (estimation risk and misspecification risk), and residual risks.

Our results suggest that, for commonly used models, misspecification risk dominates estimation risk. The analysis indicates that the multiplication factor set by the BIS is conservative if it would only be intended to cover model risk. We find, based on the standard 95%, and 99% confidence intervals, that a multiplication factor of about 2,

and 2.3, respectively would suffice for a 99% value at risk, while a multiplication factor of about 1.7 and 4, respectively would suffice for a 95% tail conditional expectation in case of the S&P 500. For the £ / $ exchange rate we find multiplication factors of about 1.6 (95% level) and 1.5 (95% level), respectively. In general the confidence levels chosen by the BIS or any other regulator need to address the trade-off between limiting the probability of excessive losses on the one hand and leaving room for operation in the market on the other hand. Besides model risk the multiplication factor set by the BIS should also cover hard-to-measure risks such as operational risk, legal risk, etc. Our framework helps disentangling the capital requirements set by the BIS into market risk, model risk, and residual risks.

Applying our methodology to hedged derivative securities we find that hedged derivative securities are much riskier than stock portfolios. For ATM straddles the capital reserve based on a 1% value at risk is about the size of the position. Splitting the model risk into estimation risk and misspecification risk, we find that the misspecification risk dominates. Investigating OTM, ATM, and ITM options we see that the Black-Scholes model performs well in modeling ITM options (which are in general easy to model). For the much harder to model OTM options we see that the misspecification risk is often above 50%.

Concluding, the framework presented allows regulators to differentiate their capital requirements on the basis of the extent to which a market can be reliably modeled on the basis of state-of-the-art technology. Depending on the performance of the model used for market risk assessment by the individual bank, model risk reserves can be determined. A further comparison between markets on the basis of the extent to which they can be reliably modeled and the determination of the size of model risk reserves for different models is left for future empirical research.

A. Risk measure derivations

A. Computation of TCE

To compute the TCE under normality we use some well-known properties of the normal and lognormal distribution. We can compute the tail conditional expectation of X when X is lognormally distributed, that is,[16]

$$\mathcal{L}\left(\log\left(X\right)\right) = \mathcal{N}\left(\mu, \sigma^2\right).$$

$$
\begin{aligned}
\text{TCE}_p\left(X\right) &= -\frac{1}{p}\int_{-\infty}^{z_p(\mu,\sigma)} \exp\left(x\right)\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) dx \\
&= -\frac{1}{p}\int_{-\infty}^{z_p(\mu,\sigma)} \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}\left(x-\left(\sigma^2+\mu\right)\right)^2 + \frac{1}{2}\sigma^2 + \mu\right) dx \\
&= -\frac{1}{p}\exp\left(\mu+\frac{1}{2}\sigma^2\right)\int_{-\infty}^{z_p(\mu,\sigma)} \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}\left(x-\left(\sigma^2+\mu\right)\right)^2\right) dx \\
&= -\frac{1}{p}\exp\left(\mu+\frac{1}{2}\sigma^2\right)\int_{-\infty}^{\frac{z_p(\mu,\sigma)-\mu-\sigma^2}{\sigma}} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y^2\right) dy \\
&= -\frac{1}{p}\exp\left(\mu+\frac{1}{2}\sigma^2\right)\int_{-\infty}^{z_p-\sigma} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y^2\right) dy \\
&= -\frac{1}{p}\exp\left(\mu+\frac{1}{2}\sigma^2\right)\Phi\left(z_p-\sigma\right) \quad (23)
\end{aligned}
$$

where $z_p\left(\mu,\sigma\right)$ denotes the p-quantile of the $\mathcal{N}\left(\mu,\sigma^2\right)$ distribution and is given by $z_p\left(\mu,\sigma\right) = z_p\sigma+\mu$, where z_p denotes the p-quantile of the standard normal distribution.

B. Asymptotic distribution of VaR and TCE

We derive the asymptotic distribution of the VaR and the TCE starting with the parametric case.

---

[16] $\mathcal{L}\left(X\right)$ denotes the law of X and $\mathcal{N}$ refers to the normal distribution.

B.1. Parametric case

We have available a data set of n (for convenience, equally spaced) returns $\left(h_1^T, ..., h_n^T\right)$ on the time interval $[0, \tau]$ which is a realization of a random sample $\left(H_1^T, ..., H_n^T\right)$, where $H_j^T \sim \mathcal{N}\left(\mu T, \sigma^2 T\right)$ for $j = 1, .., n$. μ denotes the yearly mean, σ^2 the yearly variance, and $\tau = nT$. It is well-known that the Central limit theorem gives

$$\sqrt{n}\left(\begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{bmatrix} - \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}\right) \xrightarrow{d} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2/T & 0 \\ 0 & 2\sigma^4 \end{bmatrix}\right) \tag{24}$$

with

$$\hat{\mu} = \frac{1}{\tau}\sum_{i=1}^{n} h_i^T$$

$$\hat{\sigma}^2 = \frac{1}{\tau}\sum_{i=1}^{n}\left(h_i^T - \hat{\mu}T\right)^2,$$

the maximum likelihood estimators for μ and σ^2, respectively.

Since the VaR and TCE are functions of μ and σ, their asymptotic distribution can be computed by applying the delta method (see, for example, Van der Vaart (1998)) to (24). We start with VaR. Let $\theta = (\mu, \sigma)$ and X_0 is the initial capital. All computations made with horizon T. For notational convenience, we write[17]

$$\Psi_{\text{VaR}} \equiv \text{VaR}_{m(\theta)}\left(\Pi_{m(\theta)}\right)$$
$$= -X_0\exp\left(z_p\sigma\sqrt{T} + \mu T\right). \tag{25}$$

Using

$$D\Psi_{\text{VaR}} = -\begin{bmatrix} T\Psi_{\text{VaR}} & z_p\sqrt{T}\Psi_{\text{VaR}} \end{bmatrix}$$

we get the asymptotic distribution of $\widehat{\text{VaR}}_{m(\theta)}\left(\Pi_{m(\theta)}\right)$

$$\sqrt{n}\left(\widehat{\text{VaR}}_{m(\theta)}\left(\Pi_{m(\theta)}\right) - \text{VaR}_{m(\theta)}\left(\Pi_{m(\theta)}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_{\text{VaR}_{m(\theta)}}\right), \tag{26}$$

[17]In the interest of readability the dependence on parameters is suppressed in the notation.

where

$$\Sigma_{\text{VaR}_{m(\theta)}} = T \Psi^2_{\text{VaR}} \left[\sigma^2 + z_p \sigma^3 \sqrt{T} + z_p^2 \sigma^2 / 2 \right]. \tag{27}$$

The worst-case VaR is computed by

$$
\begin{aligned}
\Psi^{wc}_{\text{VaR}} &\equiv \text{VaR}^{wc}_{m(\theta)} \left(\Pi_{m(\theta)} \right) \\
&= \text{VaR}_{m(\theta)} \left(\Pi_{m(\theta)} \right) - z_{\alpha/2} \sqrt{\Sigma_{\text{VaR}}/n}
\end{aligned} \tag{28}
$$

Using

$$D\Psi^{wc}_{\text{VaR}} = \left[\begin{array}{cc} \frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \mu} & \frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \sigma} \end{array} \right]$$

with

$$
\begin{aligned}
\frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \mu} &= -T\Psi_{\text{VaR}} \left(1 + z_{\alpha/2} \frac{\sqrt{\Sigma_{\text{VaR}}/n}}{\Psi_{\text{VaR}}} \right) \\
&= -T\Psi^{wc}_{\text{VaR}} \\
\frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \sigma} &= -z_p \sqrt{T} \Psi^{wc}_{\text{VaR}} - \frac{z_{\alpha/2}}{2\sqrt{n\Sigma_{\text{VaR}}}} \left(\sigma T \Psi^2_{\text{VaR}} \left[2 + 3z_p \sigma \sqrt{T} + z_p^2 \right] \right)
\end{aligned}
$$

we get the asymptotic distribution of $\widehat{\text{VaR}}^{wc}_{m(\theta)} \left(\Pi_{m(\theta)} \right)$

$$\sqrt{n} \left(\widehat{\text{VaR}}^{wc}_{m(\theta)} \left(\Pi_{m(\theta)} \right) - \text{VaR}^{wc}_{m(\theta)} \left(\Pi_{m(\theta)} \right) \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma^{wc}_{\text{VaR}} \right), \tag{29}$$

where

$$\Sigma^{wc}_{\text{VaR}_{m(\theta)}} = \left(\frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \mu} \right)^2 \frac{\sigma^2}{T} + \frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \mu} \frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \sigma} \sigma^3 + \frac{\sigma^4}{4} \left(\frac{\partial \Psi^{wc}_{\text{VaR}}}{\partial \sigma} \right)^2. \tag{30}$$

For the TCE we have

$$
\begin{aligned}
\Psi_{\text{TCE}} &\equiv \text{TCE}_{m(\theta)} \left(\Pi_{m(\theta)} \right) \\
&= -\frac{X_0}{p} \exp \left(\mu T + \frac{1}{2}\sigma^2 T \right) \Phi \left(z_p - \sigma\sqrt{T} \right).
\end{aligned} \tag{31}
$$

The Jacobian is given by

$$D\Psi_{\text{TCE}} = - \left[\begin{array}{cc} T\Psi_{\text{TCE}} & \left(T\sigma - \sqrt{T}\frac{\phi}{\Phi} \left(z_p - \sigma\sqrt{T} \right) \right) \Psi_{\text{TCE}} \end{array} \right],$$

45

where

$$\phi\left(\cdot\right) = \Phi'\left(\cdot\right).$$

The asymptotic distribution of $\widehat{\text{TCE}}_{m(\theta)}\left(\Pi_{m(\theta)}\right)$ is then given by

$$\sqrt{n}\left(\widehat{\text{TCE}}_{m(\theta)}\left(\Pi_{m(\theta)}\right) - \text{TCE}_{m(\theta)}\left(\Pi_{m(\theta)}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_{\text{TCE}_{m(\theta)}}\right), \qquad (32)$$

where

$$\Sigma_{\text{TCE}_{m(\theta)}} = T\Psi_{\text{TCE}}^2 \left[\sigma^2 + \left(T\sigma - \sqrt{T}\frac{\phi}{\Phi}\left(z_p - \sigma\sqrt{T}\right)\right)\sigma^3 \right.$$
$$\left. + \left(\sqrt{T}\sigma - \frac{\phi}{\Phi}\left(z_p - \sigma\sqrt{T}\right)\right)^2 \sigma^2/2\right]. \quad (33)$$

The worst-case TCE is given by

$$\begin{aligned}
\Psi_{\text{TCE}}^{wc} &\equiv \text{TCE}_{m(\theta)}^{wc}\left(\Pi_{m(\theta)}\right) \\
&= -\frac{X_0}{p}\exp\left(\mu T + \frac{1}{2}\sigma^2 T\right)\Phi\left(z_p - \sigma\sqrt{T}\right) - z_{\alpha/2}\sqrt{\Sigma_{\text{TCE}_{m(\theta)}}/n}. \quad (34)
\end{aligned}$$

The Jacobian is given by

$$D\Psi_{\text{TCE}}^{wc} = \left[\begin{array}{cc} \frac{\partial \Psi_{\text{TCE}}^{wc}}{\partial \mu} & \frac{\partial \Psi_{\text{TCE}}^{wc}}{\partial \sigma} \end{array}\right]$$

with

$$
\frac{\partial \psi_{\text{TCE}}^{wc}}{\partial \mu} = -T\psi_{\text{TCE}} \left(1 + z_{\alpha/2} \frac{\sqrt{\Sigma_{\text{TCE}}/n}}{\psi_{\text{TCE}}} \right)
$$

$$
= -T\psi_{\text{TCE}}^{wc}
$$

$$
\frac{\partial \psi_{\text{TCE}}^{wc}}{\partial \sigma} = -\left(T\sigma - \sqrt{T} \frac{\phi}{\Phi} \left(z_p - \sigma\sqrt{T} \right) \right) \psi_{\text{TCE}}^{wc}
$$

$$
- \frac{z_{\alpha/2}}{2\sqrt{n\Sigma_{\text{TCE}}}} T\psi_{\text{TCE}}^2 (\mu, \sigma) \left[2\sigma \right.
$$

$$
- \sigma^3 \left(4T + \frac{\partial (\phi/\Phi)}{\partial \sigma} \right)
$$

$$
- 3 \left(T\sigma - \sqrt{T} \frac{\phi}{\Phi} \left(z_p - \sigma\sqrt{T} \right) \right) \sigma^2
$$

$$
- \sigma^2 \left(\sqrt{T}\sigma - \frac{\phi}{\Phi} \left(z_p - \sigma\sqrt{T} \right) \right)
$$

$$
* \left(\sqrt{T} - \frac{\left(z_p - \sigma\sqrt{T} \right) \phi \left(z_p - \sigma\sqrt{T} \right)}{\Phi \left(z_p - \sigma\sqrt{T} \right)} + \frac{\phi^2}{\Phi^2} \left(z_p - \sigma\sqrt{T} \right) \right)
$$

$$
\left. - \sigma \left(\sqrt{T}\sigma - \frac{\phi}{\Phi} \left(z_p - \sigma\sqrt{T} \right) \right) \right].
$$

The asymptotic distribution is then given by

$$
\sqrt{n} \left(\left(\widehat{\text{TCE}}_{m(\theta)}^{wc} \left(\Pi_{m(\theta)} \right) - \text{TCE}_{m(\theta)}^{wc} \left(\Pi_{m(\theta)} \right) \right) \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{\text{TCE}_{m(\theta)}}^{wc} \right), \qquad (35)
$$

where

$$
\Sigma_{\text{TCE}_{m(\theta)}}^{wc} = T\psi_{\text{TCE}}^2 (\mu, \sigma) \left[\sigma^2 \left(1 + \sigma^2 T/2 \right) + \left(T\sigma - \sqrt{T} \frac{\phi}{\Phi} \left(z_p - \sigma\sqrt{T} \right) \right) \sigma^3 \right.
$$

$$
\left. + \left(\sqrt{T}\sigma - \frac{\phi}{\Phi} \left(z_p - \sigma\sqrt{T} \right) \right)^2 \sigma^2/2 \right]. \quad (36)
$$

B.2. Nonparametric case

We have available a data set of n (equally spaced) returns $\left(h_1^T, ..., h_n^T \right)$ on the interval $[0, \tau]$ which is a realization of a random sample $\left(H_1^T, ..., H_n^T \right)$, $\tau = nT$. The empirical

distribution function (EDF) is given by

$$F_n\left(y\right) \equiv \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty,y]}\left(H_i^T\right) \tag{37}$$

We have

$$\sqrt{n}\left(F_n\left(y\right) - F\left(y\right)\right) \xrightarrow{d} \mathcal{N}\left(0, F\left(y\right)\left(1 - F\left(y\right)\right)\right) \tag{38}$$

To compute the asymptotic distributions of the VaR and the TCE we need to compute the influence functions[18] of the VaR and the TCE. The value-at-risk[19] is given by

$$
\begin{aligned}
\Psi_{\text{VaR}}\left(F\right) &\equiv \text{VaR}_{m_n}\left(\Pi_{m_n}\right) \\
&= F^{-1}\left(p\right),
\end{aligned}
\tag{39}
$$

and its influence function by

$$\psi_{\text{VaR}}\left(F\right) = \frac{p - \mathbf{I}_{[y,\infty)}\left(F^{-1}\left(p\right)\right)}{f\left(F^{-1}\left(p\right)\right)}. \tag{40}$$

We can now compute the asymptotic distribution of $\widehat{\text{VaR}}_{m_n}\left(\Pi_{m_n}\right) = F_n^{-1}\left(p\right)$ as

$$\sqrt{n}\left(\widehat{\text{VaR}}_{m_n}\left(\Pi_{m_n}\right) - \text{VaR}_{m_n}\left(\Pi_{m_n}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{p\left(1-p\right)}{f^2\left(F^{-1}\left(p\right)\right)}\right) \tag{41}$$

Based on the asymptotic distribution given in (41) we can construct a confidence interval

---

[18]The influence function of Ψ can be computed as the ordinary derivative

$$\psi\left(F\right) = \frac{d}{dt}\Big|_{t=0}\psi\left(\left(1 - t\right)F + t\delta_x\right),$$

where δ_x denotes the Dirac measure.

[19]The quantile function of CDF F is the generalized inverse $F^{-1} : (0,1) \to \mathbb{R}$ given by

$$F^{-1}\left(\alpha\right) = \inf\left\{x : F\left(x\right) \geq \alpha\right\}$$

for $F^{-1}(p)$, namely

$$CI_{1-\alpha}\left(\text{VaR}_{m_n}(\Pi_{m_n})\right) = \left[\widehat{\text{VaR}}_{m_n}(\Pi_{m_n}) \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{nf^2(F^{-1}(p))}}\right], \qquad (42)$$

where z_α denotes the α-quantile of the standard normal distribution. The density f in (42) can be estimated by the Rosenblatt-Parzen kernel estimator

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right). \qquad (43)$$

The worst-case VaR is given by

$$\begin{aligned}
\Psi_{\text{VaR}}^{wc}(F) &= \text{VaR}_{m_n}^{wc}(\Pi_{m_n}) \\
&= \text{VaR}_{m_n}(\Pi_{m_n}) - z_{\alpha/2}\sqrt{\frac{p(1-p)}{nf^2(F^{-1}(p))}}. \qquad (44)
\end{aligned}$$

The influence function of the worst-case VaR is then given by

$$\psi_{\text{VaR}}^{wc}(F) = -\psi_{\text{VaR}}(F)\left[1 - z_{\alpha/2}\frac{\sqrt{p(1-p)}}{n}\frac{f'}{f}\left(F^{-1}(p)\right)\right], \qquad (45)$$

where f' can be estimated by

$$\hat{f}'(x) = \frac{-1}{nh^2}\sum_{i=1}^{n} K'\left(\frac{x_i - x}{h}\right).$$

The asymptotic distribution is given by

$$\sqrt{n}\left(\widehat{\text{VaR}}_{m_n}^{wc}(\Pi_{m_n}) - \text{VaR}_{m_n}^{wc}(\Pi_{m_n})\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_{\text{VaR}_{m_n}}^{wc}\right), \qquad (46)$$

with

$$\Sigma_{\text{VaR}_{m_n}}^{wc} = \mathbb{E}\psi_{\text{VaR}}^{wc2}(F).$$

The Tail conditional expectation is given by

$$
\begin{aligned}
\Psi_{\text{TCE}}(F) &= \text{TCE}_{m_n}(\Pi_{m_n}) \\
&= \mathbb{E}_F\left(Y|Y \leq F^{-1}(p)\right),
\end{aligned} \tag{47}
$$

and its influence function is given by

$$
\begin{aligned}
\psi_{\text{TCE}}(F) &= \frac{1}{p^2}\left[p\left(y - F^{-1}(p)\right)\mathbf{I}_{(-\infty, F^{-1}(p)]}(y)\right] \\
&= \frac{1}{p}\left(y - F^{-1}(p)\right)\mathbf{I}_{(-\infty, F^{-1}(p)]}(y) - \Psi_{\text{TCE}}(F) - F^{-1}(p). \tag{48}
\end{aligned}
$$

The asymptotic variance of $\widehat{\text{TCE}}_{m_n}(\Pi_{m_n}) = \mathbb{E}_{F_n}\left(Y|Y \leq F^{-1}(p)\right)$ is given by

$$
\begin{aligned}
\Sigma_{\text{TCE}_{m_n}} &= \mathbb{E}\psi_{\text{TCE}}^2(F) \\
&= \frac{1}{p}\mathbb{E}\left[Y^2|Y \leq F^{-1}(p)\right] - \Psi_{\text{TCE}}^2(F) \\
&\quad + \left(3 - \frac{1}{p}\right)F^{-2}(p) + \left(2 - \frac{2}{p}\right)\Psi_{\text{TCE}}(F)F^{-1}(p).
\end{aligned}
$$

The asymptotic distribution of $\widehat{\text{TCE}}_{m_n}(\Pi_{m_n}) = \mathbb{E}_{F_n}\left(Y|Y \leq F^{-1}(p)\right)$ is then given by

$$
\sqrt{n}\left(\widehat{\text{TCE}}_{m_n}(\Pi_{m_n}) - \text{TCE}_{m_n}(\Pi_{m_n})\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_{\text{TCE}_{m_n}}\right). \tag{49}
$$

A confidence interval for $\text{TCE}_{m_n}(\Pi_{m_n}) = \mathbb{E}\left(Y|Y \leq F^{-1}(p)\right)$ can be constructed using (49), namely

$$
CI_{1-\alpha}(\text{TCE}_{m_n}(\Pi_{m_n})) = \left[\widehat{\text{TCE}}_{m_n}(\Pi_{m_n}) \pm z_{\alpha/2}\sqrt{\frac{1}{n}\Sigma_{\text{TCE}_{m_n}}}\right]. \tag{50}
$$

The worst case tail conditional expectation is given by

$$
\begin{aligned}
\Psi_{\text{TCE}}^{wc}(F) &= \text{TCE}_{m_n}^{wc}(\Pi_{m_n}) \\
&= \text{TCE}_{m_n}(\Pi_{m_n}) - z_{\alpha/2}\sqrt{\frac{1}{n}\Sigma_{\text{TCE}_{m_n}}}, \tag{51}
\end{aligned}
$$

and its influence function is given by

$$
\begin{aligned}
\psi_{\text{TCE}}^{wc}\left(F\right) = & -\psi_{\text{TCE}}\left(F\right) - \frac{z_{\alpha/2}}{2\sqrt{n\mathbb{E}\psi_{\text{TCE}}^2\left(F\right)}} * \\
& \left[\frac{1}{p^2}\left(y^2 - F^{-1}\left(p\right)\right)\mathbf{I}_{\left(-\infty,F^{-1}\left(p\right)\right]}\left(y\right) - \mathbb{E}\left[Y^2|Y \leq F^{-1}\left(p\right)\right]/p - F^{-1}\left(p\right)/p \right. \\
& -2\Psi_{\text{TCE}}\left(F\right)\psi_{\text{TCE}}\left(F\right) + 2\left(3 - \frac{1}{p}\right)F^{-1}\left(p\right)\psi_{\text{VaR}}\left(F\right) \\
& \left. + \left(2 - \frac{2}{p}\right)\left(\psi_{\text{TCE}}\left(F\right)F^{-1}\left(p\right) + \Psi_{\text{TCE}}\left(F\right)\psi_{\text{VaR}}\left(F\right)\right)\right]. \quad (52)
\end{aligned}
$$

The asymptotic distribution of $\widehat{\text{TCE}}_{m_n}^{wc}\left(\Pi_{m_n}\right)$ is then given by

$$
\sqrt{n}\left(\widehat{\text{TCE}}_{m_n}^{wc}\left(\Pi_{m_n}\right) - \text{TCE}_{m_n}^{wc}\left(\Pi_{m_n}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_{\text{TCE}_{m_n}}^{wc}\right), \quad (53)
$$

where

$$
\Sigma_{\text{TCE}_{m_n}}^{wc} = \mathbb{E}\psi_{\text{TCE}}^{wc^2}\left(F\right).
$$

B. Model risk for popular risk measures

In this appendix we illustrate the model risk measure for coherent risk measures and, in particular, the worst conditional expectation and SPAN.

A coherent risk measure method ρ_m for model $m = \left(\Omega, \mathcal{F}, \mathbb{P}\right)$ can be written in the form[20]

$$
\rho_m\left(\Pi\right) = \sup_{\mathbb{Q} \in \mathcal{P}\left(m\right)} \mathbb{E}_{\mathbb{Q}}\left[\Pi\right].
$$

Different choices of $\mathcal{P}\left(m\right)$ produce different risk measures. We specify $\mathcal{P}\left(m\right)$ for WCE and SPAN.

Example 3 (WCE) Given is a model m with a base probability $\mathbb{P}$, $m = \left(\Omega, \mathcal{F}, \mathbb{P}\right)$. The

---

[20]For simplicity we use the definition given by Artzner, Delbaen, Eber, and Heath (1999). The definition for general probability spaces is given in Delbaen (2000).

class of models $\mathcal{P}(m)$ is given by

$$\mathcal{P}_{\text{WCE}}(m) = \{\mathbb{P}(.|A)\,|\mathbb{P}(A) > \alpha\}.$$

Example 4 (SPAN) Given is a model m with a base probability $\mathbb{P}$, $m = (\Omega, \mathcal{F}, \mathbb{P})$, with $\Omega = \omega_1, ..., \omega_k$ and $\mathcal{F} = 2^\Omega$, where ω_i denotes a scenario. Let

$$\mathbb{P}(\omega_1) = ... = \mathbb{P}(\omega_k) = \frac{1}{k}.$$

The SPAN method is such that [21]

$$\mathcal{P}_{\text{SPAN}}(m) \subset \{\mathbb{Q}|\mathbb{Q} << \mathbb{P}\}.$$

Note the difference between $\mathcal{P}(m)$ and $\mathcal{K}$ in Def. 11. $\mathcal{P}(m)$ is a set of probability measures based on one base probability measure $\mathbb{P}$ to compute a coherent market risk measure. However, $\mathcal{K}$ denotes a set of models. The models in this set can have different measurable spaces and different base probability measures. The model risk measure for a general coherent risk measure is then given by

$$\phi_{\text{RMM}}(\Pi, m, \mathcal{K}) = \sup_{k \in \mathcal{K}} \sup_{\mathbb{Q} \in \mathcal{P}(k)} \mathbb{E}_{\mathbb{Q}}[\Pi] - \sup_{\mathbb{Q} \in \mathcal{P}(m)} \mathbb{E}_{\mathbb{Q}}[\Pi].$$

C. Data

This appendix describes the data used in the study. In Section IV we use the total return series from DATASTREAM code: S&PCOMP(RI). In Section V we use the original S&P 500 series code: S&PCOMP. The realized dividend yield is constructed as the difference in returns of S&PCOMP(RI) and S&PCOMP. The expected dividend yield is the series given by DATASTREAM code: S&PCOMP(DY). The £/$ exchange rates is given by DATASTREAM code: USBRITP(ER). For the US risk free interest rate we have transformed the DATASTREAM series ECUSD3M(IR) to continuously

---

[21]Of course, any probability measure $\mathbb{P}^*$ equivalent to $\mathbb{P}$ could serve as a base probability measure for $\mathcal{P}_{\text{SPAN}}$. (See SPAN (1995) for details or Artzner, Delbaen, Eber, and Heath (1999) for a summary).

compounded interest rates. For the UK risk free interest rates we use the continuously compounded interest rates of the DATASTREAM series ECUKP3M(IR).

References

Artzner, Philippe, 1999, Application of Coherent Risk Measures to Capital Requirements in Insurance, *North American Actuarial Journal* 3, 11–26.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath, 1997, Thinking Coherently, *RISK* 10, 68–71.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath, 1999, Coherent Measures of Risk, *Mathematical Finance* 9, 203–228.

Basle Committee on Banking Supervision, 1996, *Amendment to the Capital Accord to Incorporate Market Risks,* Bank for International Settlements, Basle.

Basle Committee on Banking Supervision, 1999, *A New Capital Adequacy Framework,* Bank for International Settlements, Basle.

Black, Fischer, and Myron Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637–659.

Bossy, Mireille, Rajna Gibson, François-Serge Lhabitant, Nathalie Pistre, and Denis Talay, 2000, Model Risk Analysis for Bond Options in a Heath-Jarrow-Morton Framework, *Working paper RiskLab.*

Bossy, Mireille, Rajna Gibson, François-Serge Lhabitant, Nathalie Pistre, Denis Talay, and Zheng Ziyu, 2000, Volatility Model Risk Measurement and Strategies Against Worst Case Volatilities, *Journal of the French Statistics Society (Forthcoming).*

Cochrane, John H., and Jesus Saa-Requejo, 2000, Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets, *Journal of Political Economy* 108, 79–119.

Delbaen, Freddy, 2000, Coherent Risk Measures on General Probability Spaces, *Working paper ETH* pp. 1–35.

Derman, Emanuel, 1996, Model Risk, *RISK* 9, 34–37.

Dowd, Kevin, 1998, *Beyond Value at Risk: The New Science of Risk Management*. (John Wiley & Sons England).

Duffie, Darrell, and Jun Pan, 1997, An Overview of Value at Risk, *Journal of Derivatives* 4, 7–49.

Eberlein, Ernst, Jan Kallsen, and Jörn Kristen, 2001, Risk Management Based on Stochastic Volatility, *FDM preprint* 72.

El Karoui, Nicole, Monique Jeanblanc-Picqué, and Steven E. Shreve, 1998, Robustness of the Black-Scholes Formula, *Mathematical Finance* 8, 93–126.

Figlewski, Stephen, 1997, Forecasting Volatility, *Financial Markets, Institutions, and Instruments* 6, 2–87.

Galai, Dan, 1977, Tests of Market Effficiency of the Chicago Board Options Exchange, *Journal of Business* 50, 167–197.

Garman, Mark B., and Steven W. Kohlhagen, 1983, Foreign Currency Option Values, *Journal of International Money and Finance* 2, 231–237.

Gibson, Rajna, François-Serge Lhabitant, Nathalie Pistre, and Denis Talay, 1999, Interest Rate Model Risk: An Overview, *Journal of Risk* 3, 37–62.

Green, T. Clifton, and Stephen Figlewski, 1999, Market Risk and Model Risk for a Financial Institution Writing Options, *Journal of Finance* 54, 1465–1499.

Hull, John, and Wulin Suo, 2001, A Methodology for Assessing Model Risk and its Application to the Implied Volatility Function Model, *Preprint University of Toronto*.

Hull, John C., 2000, *Options, Futures, & Other Derivatives*. (Prentice Hall Toronto).

Jaschke, Stefan, and Uwe Küchler, 2001, Coherent Risk Measures and Good-Deal Bounds, *Finance and Stochastics* 5, 181–200.

Magnus, Jan, and Heinz Neudecker, 1999, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. (John Wiley & Sons Chichester, UK) revised edn.

Merton, Robert C., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4, 141–183.

Merton, Robert C., Myron S. Scholes, and Mathew L. Gladstein, 1978, The Returns and Risk of Alternative Call Option Portfolio Investment Strategies, *Journal of Business* 51, 183–242.

Merton, Robert C., Myron S. Scholes, and Mathew L. Gladstein, 1982, The Returns and Risks Characteristics of Alternative Put Option Strategies, *Journal of Business* 55, 1–56.

Newey, Whitney, and Kenneth West, 1987, A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

RISK Magazine, 1996, Value at Risk, *Risk Magazine Special Supplement* pp. 68–71.

Shorack, Galen R., and Jon A. Wellner, 1986, *Empirical Processes with Applications to Statistics*. (John Wiley & Sons New York).

SPAN, 1995, *Standard Portfolio Analysis of Risk*. (Chicago Mercantile Exchange Chicago).

Talay, Denis, and Zheng Ziyu, 2000, Model Risk Management Against Worst Case Volatility Processes for Discount Bond Options, *Working paper RiskLab*.

Van der Vaart, Aad, 1998, *Asymptotic Statistics*. (Cambridge University Press).