BANKS, INTERNAL MODELS, AND THE PROBLEM OF ADVERSE SELECTION

Christian Ewerhart[1]

Department of Economics

University of Mannheim, Germany

April 2002

Abstract. Recent empirical studies indicate that financial institutions tend to *over*estimate their market risks in disclosures to supervisory authorities. The evidence is surprising because overstated risk figures imply a costly restriction to the bank's trading activity. This paper offers a stylized model of the regulatory process, in which conservative risk reporting is the consequence of an adverse selection problem between bank and regulator. The analysis suggests that efficiency gains may be feasible when regulators disencourage overly conservative reporting, especially for banks with a history of low capital ratios. The basic argument applies also to the internal ratings-based approach to credit risk supervision. The results have an immediate bearing on the recently proposed Basel II framework.

Keywords: Bank regulation, risk management, adverse selection

JEL classification: G18, D82

[1]Postal address: Sonderforschungsbereich 504, L13, 15, D-68131 Mannheim, Germany. E-mail: ewerhart@sfb504.uni-mannheim.de. This work was presented in Madgeburg and Mannheim. For helpful comments and discussions, I thank Axel Börsch-Supan, Anette Boom, Jürgen Eichberger, Thomas Langer, Andreas Löffler, André Lucas, Benny Moldovanu, Patrick Schmitz, Deszö Szalay, Martin Weber, and Heiko Zuchel. Financial support by the Österreichische Postsparkasse and the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

1. Introduction

One of the central objectives of the prudential regulation of banks is to limit the default probability of an individual institution to a minimum possible level. Persueing this objective stabilizes the banking sector as a whole and circumvents enormous problems of measuring the social benefits generated by the banking industry.

When the return distribution of a financial institution is common knowledge, then the probability of default can be limited by requiring a minimum equity basis. Indeed, if this equity basis is chosen sufficiently large, it serves as a "cushion" and saves the bank from being liquidated in many cases. Unfortunately, however, regulators have typically only restricted information about the return-risk implications of investment strategies chosen by any individual bank. The determination of the appropriate equity basis has therefore been an issue that earned much attention among academics, regulators and representatives of the banking industry.[2]

One way to come up with a figure for the capital charge against market risk exposure is to require an equity base in terms of a percentage of the volume in which a specific asset is held, which is the so-called Cooke ratio. However, theoretical work by Koehn and Santomero (1980) revealed that, if this regulatory ratio is uniform over a larger class of assets with heterogeneous return-risk structure, then tigher regulation may be detrimental for the stability of the individual bank and therefore potentially for the whole financial sector. Specifically, it was shown that lowering the capital ratio may induce the financial institution to reshuf-

[2]See, e.g., the special issue of the Journal of Banking and Finance beginning with an article by Berger, Herring, and Szegö (1995). For a more general survey on banking regulation, see Bhattacharya, Boot, and Thakor (1998).

fle its portfolio into more risky investments, thereby potentially increasing its probability of default.

A somewhat more refined approach uses specific ratios for assets of different risk classes. Indeed, as Kim and Santomero (1988) show, if these ratios are set to their theoretically correct values, then one can obtain an effective instrument that limits the default probability and simultaneously induces mean-variance efficient investment. Paralleling these theoretical developments, in 1988, a transnational working group on banking supervision, the Basel Committee, proposed and implemented worldwide the so-called Basel Capital Accord, which implemented some of the ideas underlying the previously mentioned theoretical approach.

While the implementation of this system was an important step towards effective prudential regulation of the financial sector, it did not solve all problems of capital adequacy. E.g., the above-mentioned theoretically correct values for the equity ratios may in practice be difficult to obtain, e.g., because the measurement of the correlations between sufficiently many risk classes may be impossible. Moreover, many modern financial instruments such as stock options and convertible bonds carry non-linear risks, while the positive result seems to depend on the assumption of normal return distributions.

Given that banks already have in place effective tools for the measurement of market risks, it therefore seemed as a good idea to ask inhouse risk management to report their numerical estimates for the banks' risks. Specially adapted statistical tests, known as *backtesting*, could then be used to assess the accuracy of the reported numbers, and an appropriate punishment could be imposed if the risk model appears to be imprecise or manipulated. Specifically, in case of understatement of risks, regulators would tighten the capital requirements for the next quarter by multiplying capital charges by a suitable factor that is increasing in the number of days at which the predicted risks are exceeded by actual losses (cf. Table 1). This concept, known as the internal models approach, was released as an amendmend to the capital accord by the Basel Committee (1996a, 1996b), and implemented from the beginning of 1998 onwards in many industrialized countries.

- place Table 1 here -

The capital accord is currently under revision, yet with the declared intention to leave the supervisiory framwork for market risks essentially unchanged. Moreover, the framework for credit risk regulation has been proposed to follow the so-called internal ratings-based approach, that is structurally similar to the internal-models approach (cf. Basel Committee, 2001). The present study argues that these approaches, unless suitably modified, may generate a profit-diminishing informational externality especially on those banks that are least likely to default.

More specifically, we argue in the paper that in order to avoid this externality, the regulator should set incentives that *disencourage conservative risk reporting.* It will be shown that if these incentives are absent, as it is the case under both the existing capital accord and the new proposals, then the regulator will not obtain any additional information from the disclosed risk data, and, more severly, prudent banks are likely to restrict profitable trading activities more than socially desirable, leaving a non-marginal fraction of their regulatory capital unexploited.

The analysis also helps to explain the somewhat surprising conservatism observable from actual risk data. To see the problem, note that intuitively, the strong leverage effect of a better equity exploitation overcompensates the limited risk of a penalty factor, which suggests that some banks may be willing to sail close to the wind concerning their risk reporting. This point has been made likewise by academics and supervisors (see the panel discussion on this topic documented by Hellwig and Staub, 1996). Indeed, it was noted that the official back-test is skewed in favor of accepting models (cf. Basel Committee, 1996c). It has also been confirmed numerically (see, Lucas, 2001, and Ewerhart, 2000), that in practice, banks should tend to understate their risks. The reason for this fact is the high profitability of risk-bearing investment banking activities on the one hand, and the regulator's limited leverage for regulatory penalties on the other, which is partly due to the coexistence of more and less refined regulatory schemes.

Contrasting these intuitions and numerical findings, recent empirical evidence suggests that, to the contrary, many relevant banks in fact overestimate their risks, and do not exploit the regulator's deliberation. We will survey the existing evidence in Section 2.

This paper proposes a model of the regulatory process, and offers an explanation of the seemingly contradicting evidence. In the model, the empirically observed effect is originated by an adverse selection problem in the relationship between supervisor and bank. As the bank cannot credibly communicate its risk exposure, and the supervisor needs to restrict the default probability of *all* banks, more prudent banks have to bear an informational externality which is realized in the choice of an overly conservative investment strategy.

The rest of the paper is structured as follows. Section 2 gives an overview over some empirical observations made in connection with the internal models approach. In Section 3, we review the Pyle-Hart-Jaffee model of portfolio choice. Section 4 introduces risk regulation on the basis of value-at-risk figures, and offers a positive result that states the effectiveness of value-at-risk regulation in a pure adverse selection setting. In Section 5, we introduce noise to the regulators observations, and show how under this assumption, our model predicts conservative investment and reporting as a consequence of an adverse selection problem. Section 6 discusses the some extensions and robustness. Section 7 concludes with a brief summary and policy implications.

2. Empirical observations

There are a number of recent studies that discuss the performance of models-based capital charges for market risks.[3] To understand the results of these studies, it is useful to briefly review two important notions commonly used in the Basel regulatory framework. For further details on the framework, we refer the reader to Basel Committee (1996c).

The first is the standard risk measure used in bank disclosures which is the so-called *value-at-risk* (VaR).[4] The value-at-risk corresponding to a 1% confidence level looks forward a specific period of time (the halting period), and is the loss to be exceeded with a 1 in 100 chance. The pervasive use of the 99% confidence level in the definition of the value-at-risk is of course a matter of convention. An alternative confidence level often used in practice, especially in J.P. Morgan's Risk Metrics framework, is 95%, where the chances become 1 in 20, and the risk figure is correspondingly lower.

The second notion is that of an *exception*. When the accuracy of a bank's internal model is assessed by bank supervison in a quarterly review, regulators use a

[3]These studies can be considered as a part of a broader program that analyzes the overall performance of capital regulation. A recent contribution in this vein is Rime (2001), who examines capital holdings and risk taking behavior by Swiss banks in response to changes in the regulatory setting (before the implementation of the amendment).

[4]Baumol (1963) proposed the use of quantiles as risk measures in the Markowitz portfolio selection approach. See Jorion (1995) for a general introduction to the value-at-risk methodology.

simple non-parametric test that counts the number of days out of the past 250 trading days at which the bank's trading activities resulted into losses that exceed the predicted values of the internal risk model for the same day. This number is referred to as the number of exceptions, and is used by the regulators to determine a factor on the capital charges for the next period, as shown in Table I, where a higher factor means a tigher restriction on trading activities.

Equipped with the above notions, we are now able to survey the empirical evidence. The first paper that examined data on financial instititions' risk disclosures is a survey by the Basel Committee (1999). Written only briefly after the implementation of the internal models approach, this study analyzed daily data of over 40 banks, located in 9 countries, in the second half of the year 1998.

The results concerning the number of exceptions are as follows. Almost half of the surveyed institutions had no exceptions within the above-mentioned period. Of those banks that experienced exceptions, the majority reported less then five exceptions. Only three banks reported five or more exception, the maximum being seven exceptions. These numbers are interpreted by the Basel Committee as providing preliminary, but otherwise strong evidence for the sufficiency of the capital charges resulting from the internal models approch.

We will rededicate this anecdotal data into a first indication that risk reporting is overly conservative. Indeed, a simple binomial test shows that *the null hypothesis that no bank overstates its true value-at-risk can be rejected at a confidence level of 99%.* (Here we assumed that the second half of 1998 had 125 trading days, that the value-at-risk is calculated at a 99% confidence level, and that returns are independent over time, which yields a probability of less than 29% that a bank would have no exceptions. Putting this into a binomial test with 40 trials and 18 or more successful outcomes gives a likelihood of 0.8%).

A second empirical paper on the internal models approach is by Berkowitz and O'Brien (2001). It covers a longer period of time, vz. 01/98 through 03/00. The number of considered institutions is six and therefore smaller than in the first study, yet these six banks include the largest US bank derivative dealers, and are all among the 10 biggest financial institutions measured in terms of notional amounts outstanding as of year-end 1999.

Among other things, Berkowitz and O'Brien (2001) point out in their study that *banks's value-at-risk estimates tend to be conservative with respect to the 99th percentile of actual P&L.* They come to this conclusion from two observations.

Firstly, they shows that for five out of six banks, the average reported value-at-risk lies outside the lower 99th percentile of the P&L statistics, with value-at-risks for four banks ranging from 1.6 to over 3 times their respective 99th percentile P&Ls.

Secondly, they observe that in the considered 500 trading days, only for one bank the actual P&L exceeded the reported risk figures more than three times, while the expected number of exceptions is five.

The authors trace the conservatism of the risk reports back to various bank model features and regulatory constraints. Specifically, they name as a partial explanation the fact that all considered banks excluded a significant component of income (specifically, net fee income) from value-at-risk estimates. While this explains the conservatism, it does not explain why banks do not account for this measurement error on average given that equity is generally considered as a scarce resource in the banking industry. After all, one could have rescaled the risks figured into more realistic domains using appropriate discount factors.

Another source of conservatism referred to is the common practice in applied risk management to sum up various subgroup value-at-risks. E.g., if an instrument carries currency as well as interest rate risks, then the value-at-risk estimates obtained for a partial risk analysis are simply added. This method neglects potential diversification effects and may therefore lead to overstated risks. While it is unlikely that financial institutions had access to significantly better models at that stage, this is again more an argument for the inaccuracy of the models, and less for the conservatism on average.

Berkowitz and O'Brien (2001) also note the problem that the one-sidedness of the "backtest" induces banks to be conservative in their estimate. However, they do not mention that a more conservative risk report is also costly because it implies a more conservative investment strategy. In fact, as mentioned before, economic analyses of the trade-off suggest that if one only considers the cost of higher capital charges vis-a-vis the benefit of better equity exploitation, then the Basel incentive structure is fairly weak and implies an *under*statement of risks rather than conservatism.

So while there are imperfections in the data and methodology, these only explain an uncertainty about risk figures, but not why there are so few exceptions *on average*.

As above, we can again use a simple binomial test in order to show that also in this case, the null hypothesis that no bank overstates its true value-at-risk can be rejected at a confidence level of 99%. (The test is based on the assumption of 500 trading days. The probability that a given bank has 3 or less exceptions in 500 days is approximately 26.4%. The probability that at most one bank has more than three exceptions is then lower than 1%, which rejects the hypothesis).

There are further indications of conservative risk reporting. Gizycki and Hereford (1998) asked 22 financial firms in Australia to calculate value-at-risk figures for a number of synthetic portfolios. They find a wide dispersion in the risk figures, and trace this observation back to conservatism in a number of models. Moreover, reconfirming our previous point that the method of risk measurement should not be made responsible for conservatism, this study also shows that there is no significant correlation between the chosen method of calculation and the risk figures.

Jorion (2001) does an extensive empirical study on the informativeness of value-at-risk disclosures by 8 large U.S. American commercial banks. The analysis uses quarterly data between 12/94 and 09/00. Jorion's results exhibit a striking contrast between the informativeness of the value-at-risk figures at the individual bank and cross-sectional levels. More specifically, he shows that for *only two* out of eight banks, the official value-at-risk disclosures are a significant predictor of trading income variability at a confidence level of 95%. Some banks even reported nearly constant value-at-risk figures over the considered period.

Cross-sectionally, however, the picture changes, and one obtains significant results in favor of informativeness of the risk figures. A methodological problem here is the size effect, which can be accounted for alternatively by an extended regression and scaling with respect to notional amounts. While the results turn out to be robust with respect to these changes in the specification, it will be noted that they do not contradict our predictions because the appropriate test would be to scale the data with respect to regulatory equity.

There are other reasons to believe that Jorion's (2001) study does not provide evidence for non-conservative disclosures to supervisory authorities. E.g., the data set begins prior to the implementation of the internal models approach. Moreover, the disclosed figures are used as predictors for unexpected trading income *in the next* rather than in the current quarter. Jorion's results could be a consequence of the facts that asset returns display volatility clustering, and that banks typically use restrictions on deliberation in addition to VaR limits in their management of individual trading units. This would render investment not fully responsive to changes in the market conditions, and higher value-at-risk figures would result in periods of higher price volatility.

From the material surveyed above, we conclude that market risk reporting by financial institutions tends to be overly conservative for a significant fraction of the banking population. In the next two sections, we offer an explanation for this observation that is based upon asymmetric information between bank and regulator.

3. Portfolio choice

Our model is based on the standard portfolio model used by Pyle (1971) and Hart and Jaffee (1974).[5]

The financial institution is endowed with equity E. There is one investment period, at the beginning of which the portfolio manager may buy or sell securities in the financial markets as described in the sequel.

- place Figure 1 here -

[5]Existing theories of banking are surveyed by Bhattacharya and Thakor (1993). See also Freixas and Rochet (1997).

There is a safe asset with return $R_0 > 0$. There are also L risky securities or assets, whose unit prices are normalized to 1. The vector of random returns of these assets

$$\widetilde{R} = (\widetilde{R}^1, ..., \widetilde{R}^L) \tag{1}$$

follows a multivariate normal distribution, with expected returns

$$E[\widetilde{R}^l] = R_0 + \rho^l, \tag{2}$$

and variance-covariance matrix - $\in \Re^{L \times L}$. We assume that no asset can be represented as a linear combination of the other securities, so that - is invertible.

There are no constraints to short-selling, and therefore a portfolio can be described by a vector

$$\alpha = (\alpha^1, ..., \alpha^L) \in \Re^L, \tag{3}$$

where α^l denotes the amount invested in the l-th security.

Total return from trading is

$$Z = (E - \sum_{l=1}^{L} \alpha^l) R_0 + \sum_{l=1}^{L} \alpha^l \widetilde{R}^l, \tag{4}$$

which may be positive or negative.

The financial institution may default. Specifically, when $E + Z$ is negative, the firm is liquidated.

The manager is assumed to have a non-decreasing and concave utility function $U(Z)$. Keeley and Furlong (1990) have noted an inconsistency in previous work on the regulation of portfolio managing banks. The point is that the bank cannot be assumed both to underly the moral hazard problem and to be of unlimited liability at the same time. Rochet (1992) has analyzed the behavior of a bank with limited liability, and shows in particular that insufficienty capitalized banks may exhibit risk-loving behavior. We will discuss in section 6 how our results are affected in the presence of risk-loving behavior. All we need for the moment is that the bank's (restricted) optimization problem is well-defined, and that the corresponding set of solutions is a connected set in the (μ, σ) plane.

The manager chooses a portfolio α as to maximize his expected utility

$$\overline{U}(\mu, \sigma) = \int U(\mu + t\sigma) d\Phi(t), \tag{5}$$

where

$$\mu = R_0 E + \alpha^T \rho \tag{6}$$

and

$$\sigma^2 = \alpha^T \text{-}\, \alpha \tag{7}$$

are mean and variance of the chosen portfolio, and $\Phi(.)$ denotes the cumulative distribution function of the standard normal distribution.

It is known in the literature that the concavity of $U(.)$ implies that the indirect utility function is concave in (μ, σ)-space, so that indifference curves are convex (cf. Rochet, 1992, Proposition 5).

Proposition 1. *The manager chooses a portfolio from the efficient frontier, which is a straight line in (μ, σ)-space with positive slope $\sqrt{\rho^T \text{-}\, ^{-1} \rho}$, and which intersects the μ-axis at $R_0 E$ (cf. Figure 1).*

Proof. The manager's problem is

$$\alpha^* = \arg\max_{\alpha} \overline{U}(\mu, \sigma) \tag{8}$$

$$\text{s.t.} \tag{9}$$

$$\mu = R_0 E + \alpha^T \rho \tag{10}$$

$$\sigma = \sqrt{\alpha^T \text{-}\ \alpha} \tag{11}$$

¿From the first-order condition,

$$\alpha = \nu \text{-}\ ^{-1} \rho, \tag{12}$$

where

$$\nu = -\sigma \frac{\partial \overline{U} / \partial \mu}{\partial \overline{U} / \partial \sigma}. \tag{13}$$

This implies

$$\mu^* = R_0 E + \nu \rho^T \text{-}\ ^{-1} \rho \tag{14}$$

$$\sigma^* = \nu \sqrt{\rho^T \text{-}\ ^{-1} \rho} \tag{15}$$

The assertion follows. □

4. Pure adverse selection

The regulator's objective is to maximize welfare under the condition that each individual financial institution defaults with a probability of at most $p_0 \in (0, 1]$. As argued in the introduction, this may be the best thing to strive for, given the interdependencies within the financial sector, and the problems in measuring consumer surplus from financial intermediation.

Given that the regulator needs to reduce the probability of default to p_0, he must make sure that the financial institution uses an investment strategy that is on the left-hand side of the iso-default-probability curve depicted in Figure 2. When the target probability of default decreases, the implied restrictions on the portfolio choice are more restrictive, and the iso-default-probability curves, anchored at the point $(0, -E)$, tilt towards the left side.

- place Figure 2 here -

Assume for the moment that the regulator observes the return-risk structure (μ, σ) that was chosen by the portfolio manager, but not his type θ. Following standard terminology in economic theory (cf., e.g., Calliaud, Guesnerie, and Rey, 1992), we refer to this setting as one of pure adverse selection. The regulation is based upon value-at-risk disclosures, all calculated with respect to a confidence level of $1 - p$, fixed once and for all. The *value-at-risk* of a portfolio (μ, σ) with respect to a confidence level $1 - p$ is defined as

$$V = -\sigma \Phi^{-1}(p) - \mu, \tag{16}$$

where $\Phi^{-1}(.)$ denotes the inverse of the cumulative distribution function of the standard normal distribution.

Figure 3 shows iso-value-at-risk curves in return-risk space. It can be seen that these curves are parallel straight lines, and that any such line intersects the μ-axis at the value $-V$, where V is the corresponding value-at-risk.

- place Figure 3 here -

Current regulatory documents do not display an explicit target probability p_0. However, the implicit target probability is approximately 1% or lower, as one can see from the recommended confidence level of $1 - p = 99\%$, and from the minimum capital multiplicator of 3 (cf. Basel Committee, 1996c).

Assumption 1. *The confidence level used for the measurement of the value-at-risk* $1 - p$ *is sufficiently large. More precisely,*

$$p < \Phi(-\sqrt{\rho^T \text{-}\ {}^{-1} \rho}) \tag{17}$$

This assumption assures that the iso-value-at-risk lines are steeper than the efficient frontier, and thereby helps to eliminate some uninteresting cases. E.g., from Figure 3, one sees that volatility and value-at-risk are positively correlated risk measures. As we will now show, another consequence is that for any given value-at-risk, there is a unique return-risk structure (μ, σ) with this value-at-risk.

Proposition 2. *Under Assumption 1, any portfolio (μ, σ) on the efficient frontier can be uniquely characterized by its value-at-risk*

$$V = -\sigma\Phi^{-1}(p) - \mu. \tag{18}$$

Proof. The set of points with the same value-at-risk V is a line given by

$$p = \Phi(\frac{-V - \mu}{\sigma}), \tag{19}$$

which crosses the μ-axis at $-V$ and which has slope $\Phi^{-1}(p)$ in the (σ, μ) plane. Under Assumption 1, there is precisely one intersection of this steeper line with the line that constitutes the efficient frontier. □

We will now introduce a more specific utility function for the manager which captures both the heterogeneity of banks with respect to their short-term risk attitude and the ex-post nature of currently practiced market risk supervision. Assume that utility is intertemporally additive and can be represented as

$$U = U_0(Z, \theta) + U_1(Z, \lambda_1), \tag{20}$$

where the first term $U_0(Z, \theta)$ represents the utility earned in the present period, and θ is referred to as the manager's short-term *risk attitude.*

The parameter θ reflects the fact that banks may be of different types. Heterogeneity may result from different reasons, even in the presence of homogeneous shareholder structures. E.g., banks may differ in their history of profitability, and their availability of hidden reserves. The necessity to provide the correct signals to the capital market may then induce banks to persue different short-term objectives. On the level of individual managers, a reason for heterogeneous short-term objectives may be differences in the horizons to either retirement or expected termination of the labor contract.

The second term $U_1(Z, \lambda_1)$ represents the utility earned from the bank charter in the next and all subsequent periods, where λ_1 is a regulatory parameter imposed by the supervisor for the next period and for possibly further periods. We assume that $U_1(Z, \lambda_1)$ is non-increasing in λ_1, so that intuitively, a higher λ_1 corresponds to a stronger punishment by the regulator.

Under current regulation, λ_1 could e.g. be the Basel plus factor, which means essentially a restriction to trading activities. Alternative interpretations are possible, e.g., λ_1 could represent the probability of an audit in the next period, or the value of λ_1 may entail the right to use an internal model rather than the standardized method in the next period.[6]

The parameter $\theta \in [\underline{\theta}, \overline{\theta}]$ is ex-ante uncertain, and assumed to be drawn at the beginning of the trading period, and according to a cumulative distribution function $F(\theta)$. The distribution of types is common knowledge. We envisage managers with high θ to be more prudent in the sense that they have a weaker inclination to risky investments. This is formalized in our next assumption.

Assumption 2. *The bank's (short-term) marginal expected disutility from risk-taking is nondecreasing with the type θ in the relevant domain. That is, on the efficient frontier,*

$$\frac{\partial^2 \overline{U}}{\partial \theta \partial V} \leq 0. \tag{21}$$

This single-crossing assumption is satisfied for a variety of utility functions. E.g., it is not hard to show that Assumption 2 is satisfied when short-term utility is

[6] In fact, the formal arguments go through if λ_1 is taken from an arbitrary partially ordered set.

quadratic

$$U_0(Z,\theta) = Z - \theta Z^2, \tag{22}$$

or CARA

$$U_0(Z,\theta) = -\exp(-\theta Z). \tag{23}$$

In both cases, U_1 can be arbitrary, as long as U remains concave in Z.

Proposition 3. *In the absense of regulation, the preferred risk exposure $V^*(\theta)$ is nonincreasing in the manager's type θ.*

Proof. Note that for a constant λ_1, indirect utility $\overline{U}(\theta, V)$ is submodular. Hence the assertion follows from Topkis' monotonicity theorem (cf. Milgrom and Roberts, 1990). □

To ensure that the default probability is limited, the regulator requires that the bank's equity holdings are sufficiently big, i.e., he requires

- place Figure 4 here -

$$E \geq \lambda_0 \widehat{V}, \tag{24}$$

for some factor λ_0 that is generally understood to take account of the limited precision of the disclosed risk figures, and which also serves as a means to penalize misreporting in the previous period, as suggested by the values in Table I.

Theorem 1. *Assume that the regulator requires that equity must cover a multiple of the portfolio's value-at-risk, i.e.,*

$$E \geq \lambda_0 V. \tag{25}$$

Then, in the pure adverse selection setting, if λ_0 is chosen appropriately, there is a risk attitude θ^* such that any manager of type $\theta > \theta^*$ chooses a portfolio of risk p_0, and any manager of type $\theta \leq \theta^*$ is unaffected by the regulation. Thus, if *the regulator can observe the return-risk structure of individual portfolios, VaR regulation is an efficient way to guarantee that any individual financial institution defaults with a probability of at most* p_0.

Proof. Let $V[p]$ denote the value-at-risk of the bank's portfolio corresponding to a confidence level of $1-p$. Let λ_0 be such that

$$\lambda_0 V[p] = V[p_0]. \tag{26}$$

Let now θ^* be the type that chooses the portfolio with default probability p_0 in the absense of regulation. Then from Proposition 3, each type $\theta \geq \theta^*$ is unaffected by regulation, which is merely a restriction on the investment possibility set. For $\theta < \theta^*$, however, the convexity of the indifference curves implies that the bank chooses a default probability of p_0. This is because if not, the indifference curve must lie on the efficient frontier, and go through the point characterized by the default probability p_0, which is a contradiction. □

Theorem 1 says that value-at-risk reporting effectuates that the bank will not default with probability higher than p_0. The idea of the proof is the following. As illustrated in Figure 4, the iso-value-at-risk curve corresponding to V is a straight line given by

$$V = -\sigma\Phi^{-1}(p) - \mu, \tag{27}$$

which crosses the mean axis at $-V$ and which has slope $\Phi^{-1}(p)$ in the (σ, μ) plane. But, from comparing Figures 2 and 3, one can see that for a suitable choice of λ_0, the iso-value-at-risk curve corresponding to a risk exposure V and confidence level $1-p$ imposes the same restriction on the efficient frontier as the iso-default-probability curve corresponding to a probability p_0, which implies that regulation is effective and non-distorting.

Note that in contrast to the standard approach regulation using capital charges for risk-weighted assets (cf. Koehn and Santomero, 1980), VaR regulation never induces banks to choose portfolios that do not lie on the efficient frontier. The standard approach may generate mean-variance inefficiency when risk weight are different from their theoretically correct values because a mean-variance improving modification of the portfolio composition may have a positive shadow price. This cannot happen under VaR regulation because a mean-variance improvement always lowers the value-at-risk, so that efficiency gains are always within regulatory deliberation.

5. Noisy adverse selection

We assume now that the institution's return-risk structure (μ, σ) cannot be credibly communicated to the regulator.

As the regulator does not know the portfolio's return-risk structure, he must set appropriate incentives for the portfolio manager (see Figure 5). The regulator therefore asks the manager for an estimate $\widehat{V}$ of the value-at-risk. Alternatively, the regulator may ask for (μ, σ), or for the default probability p.

- place Figure 5 here -

We assume however that the regulator has a *backtesting technology* at hand that generates a random signal k whose distribution depends on the actual value-at-risk V and on the reported $\widehat{V}$.

The parameter k is a measure of how conservative or aggressive the reported risks $\widehat{V}$ are when compared to the true risks V. A high k corresponds to probably understated risks, a low k to probably overstated risk. Under current regulatory practice, the parameter k is discrete and corresponds to the number of exceptions in the Basel Committee (1996c) backtesting procedure.

Denote by $G(k|V,\widehat{V})$ the cumulative distribution function of the indicator variable k given true risks V and reported risks $\widehat{V}$. Our next assumption says that the number of exceptions tends to be higher if either the true risks are higher or the reported risks are lower. Mathematically, this is captured in terms of first-order stochastic dominance as follows:

Assumption 3. *The cumulative distribution function $G(k|V,\widehat{V})$ is non-decreasing in V, and non-increasing in $\widehat{V}$.*

The first part of the assumption captures the regulatory costs from high-risk investment. If V is chosen very high, then, as the reported value-at-risk is bounded by the equity constraint, the number of exceptions is large with a higher probability, and on average a stricter penalty results.

We assume that it is regulatory policy to persue the backtesting procedure, to produce the k, and to impose a penalty parameter

$$\lambda(k) \in [\underline{\lambda}, \overline{\lambda}], \tag{28}$$

which is assumed to be strictly increasing in k.

Assume that the regulator has designed the incentive system in a way that induces a bank of risk type θ to choose an actual value-at-risk $V(\theta)$, and to report a hypothetical value-at-risk $\widehat{V}(\theta)$.

Theorem 2. *Assume that the regulation affects some type of financial institution. Then, all but the most risk-prone type $\underline{\theta}$ use inefficiently conservative investment strategies that leave regulatory risk capital unexploited. Moreover, all banks report maximum risks*

$$\widehat{V}(\theta) = \widehat{V}(\underline{\theta}) = E. \tag{29}$$

Proof. By Proposition 2, there is a one-to-one relationship between efficient portfolios (μ, σ) and corresponding value-at-risk figures $V \in [0, \infty)$. Write $\mu(V)$ and $\sigma(V)$ for the mean and standard deviation of the efficient portfolio with value-at-risk V.

Let

$$\tau(V, \widehat{V}) = \int \int U_1(\mu(V) + t\sigma(V), \lambda(k)) dG(k|V, \widehat{V}) dt \tag{30}$$

be the expected net present value of a bank charter if today's actual value-at-risk is V and today's reported value-at-risk is $\widehat{V}$.

Let $V(\theta)$ and $\widehat{V}(\theta)$ denote the equilibrium choices of a manager of type θ of actual and reported risks, respectively. To induce the manager of type θ to choose actual risk exposure $V(\theta)$ and to report risks $\widehat{V}(\theta)$, the incentive compatibility condition

$$\overline{U}_0(V(\theta), \theta) + \tau(V(\theta), \widehat{V}(\theta)) \geq \overline{U}_0(V, \theta) + \tau(V, \widehat{V}) \tag{31}$$

must be satisfied for all pairs $(V, \widehat{V})$, where $\widehat{V} \leq E$. It is clear from Assumption 3, and because $U_1(Z, \lambda)$ is non-increasing in λ that $\tau(V, \widehat{V})$ is non-decreasing in $\widehat{V}$. Hence all portfolio managers will report $\widehat{V} = E$, which proves the second part of the assertion. Write now $\underline{\tau}(V) := \tau(V, E)$. Then (31) simplifies into

$$\overline{U}_0(V(\theta), \theta) + \underline{\tau}(V(\theta)) \geq \overline{U}_0(V, \theta) + \underline{\tau}(V) \tag{32}$$

¿From Assumption 1, we know that $\overline{U}_0(V, \theta)$ is submodular. Hence also $\overline{U}_0(V, \theta) + \underline{\tau}(V)$ is submodular. But then, it follows from Topkis' monotonicity theorem that $V(\theta) < V(\overline{\theta}) = E$ for all $\theta < \overline{\theta}$. □

Note that Theorem 2 does not say that investments are mean-variance inefficient. The inefficiency manifests itself in that, under VaR regulation and noisy information about the return-risk consequences of the bank's trading activities, prudent banks will in the presence of more risk-prone banks choose portfolios that are overly conservative, and consequently earn lower expected profits.

The consequences of Theorem 2 are striking: When the safe asset is not sufficiently profitable, then the informational externality may make portfolio management altogether unprofitable in comparision to respect to some existing outside option, so that all sufficiently prudent firms will leave the market.

While our presentation focusses on market risks, it should be clear that the basic arguments apply with minor modification to the internal ratings approach to *credit risk*, as favored by the Basel Committee (2001) in the Basel II revision of the Capital Accord. In fact, the problem is probably much more severe for credit risk supervision because of the time lag between credit approval and repayment.

6. Discussion

This section serves three purposes. First, it evaluates some alternative explanations of conservative risk reports. Then, a number of extensions of the model are considered, and robustness is discussed. Finally, we touch upon the question of the optimal contract.

Alternative explanations. We perceive momentarily two alternative explanations for conservative risk reporting. The first is the existence of *informational frictions between capital markets and bank management.* Under this condition, the prudent manager may want to use the backtesting technology to credibly convey information about his "quality" to the market. Note, however, that given the ex-post nature of the backtest, signaling will be relevant only when the financial institution has some relevant long-term characteristic. Moreover, it is not so clear whether the backtesting technology could serve as a signaling device in the first place. This is because the disclosures alone can convey information only about the relative costs resulting from the supervisor's penalty scheme, which need not necessarily be correlated with the bank's long-term characteristic that the investor is interested in.

Another posssible explanation is the political economy within the bank. Specifically, from a practical perspective, limited exploitation of the regulatory capital may result from the need to subdivide the regulatory capital for the whole financial institution into *limits for individual traders.* As it is virtually impossible to coordinate the trading activities between these traders in a way that correlations are taken account of, the division of the regulatory capital is typically rather rigid. Then, even if all traders use up their limit almost completely, the diversification effects are likely to lower the resulting value-at-risk for the whole firm, so that some regulatory capital would remain unexploited. However, current regulation calculates the capital charge for market risk essentially as an average of the value-at-risks over the last 60 trading days. More precisely, the requirement can be stated as

$$E_t \geq \max\{V_{t-1}, \frac{\lambda_0}{60} \sum_{\tau=1}^{60} V_{t-\tau}\} \tag{33}$$

In practive, $\lambda_0 \geq 3$, which allows a bank to "tune" its usage of the regulatory capital by averaging over time, even if averaging over individual traders may not be possible.

Extensions and robustness. The arguments presented are typically robust with respect to straightforward extensions. To start with, *investment opportunities may be restricted.* E.g., while we assumed the existence of a safe asset, there seems to be no need for that. The only complication that may arise in this case is that, because the efficient frontier becomes a hyperbola, the geometry of the VaR-regulation becomes slighly more complex. The key arguments, however, are unaffected. A similar remark applies to *short-sales constraints*, which also mean only a deformation of the efficient frontier.

We have assumed throughout that equity is fixed, and cannot be increased easily. This is of course a simplification because many financial institutions will be able to *enlarge the capital base*, e.g., by making seasoned offerings. Note, however, that the request for further equity is often interpreted by the providers of capital as a signal for weaker future performance, so that such an request will not always be in the bank's interest.

When *the bank may be risk-loving*, then our results are still valid as long as the bank's restricted optimization problem is well-defined and has a unique solution. When the indifference curve touches the efficient frontier more than once, then the VaR restriction may induce the bank to switch discontinuously to some point on the efficient frontier strictly below the VaR constraint. Thus, the literal statement of Theorem 1 would cease to hold, but as long as we stay in the pure adverse selection setting, capital adequacy using VaR would remain an effective and efficient instrument of market risk supervision. A similar remark applies to Theorem 2.

In the current setting, the regulator has lexicographic preferences, where the limitation of failure risk is the first objective, and having banks making profits is the second. In a more realistic picture, the regulator would have a welfare function that incorporates the *tradeoff between raising the bank's surplus and incurring additional risks for the society.* It is not difficult to see, however, that in focussing on the risks of an investment, current VaR regulation will be too inflexible to take account of this tradeoff, and will therefore be inefficient even in the pure adverse selection setting.

The optimal mechanism. The optimal mechanism for the regulator will make the penalty dependent not only on the outcome of the backtest, but also on $\widehat{V}$.[7] This makes a difference if, for a given value-at-risk V, the bank's expected utility realized in the next and all subsequent periods is *not* monotonic in $\widehat{V}$. Note, however, that the mathematical derivation of an optimal regulatory scheme would make it necessary to write out the utility function

$$U = \sum_{t=0}^{\infty} \delta^t \widetilde{U}_t(Z_t, \theta_t, \lambda_t) \tag{34}$$

as a sum of future utility components, where $U_0 \equiv \widetilde{U}_0$. This is because the regulators penalty scheme $\lambda(.)$ affects not only next period's investment possibility set, but also the bank's objective function for the next period. In this sense, the U_1 in the basic model is properly speaking an indirect utility function that already incorporates how $\lambda(.)$ affects future utility components $\widetilde{U}_t$ for $t \geq 1$. It would therefore be inconsistent to optimize $\lambda(.)$ in the previously developed model for a given function U_1.[8]

The above analysis nevertheless suggests a potential lever for obtaining efficiency gains in regulation based on internal risk assessment. As the adverse selection problem originates from the supervisor's inability to tell apart banks of heterogeneous types on the basis of their value-at-risk reports, which again is due to the monotonicity of the incentive scheme, it seems that a non-monotonic scheme

[7] According to the *taxation principle*, asking for the bank's type θ does not convey additional information (cf. Caillaud, Guesnerie, and Rey, 1992).

[8] Notwithstanding the technical difficulties that are likely to appear in the derivation of the optimal mechanism, this issue seems very relevant, and we hope to be able to address it in future work.

could help to mitigate the problem. Specifically, it appears that the optimal incentive scheme should not only penalize risk figures that are too low, but also disencourage conservative risk reporting. Intuitively, this will make sure that the regulator learns something from the reported risk figures about the actual risks of the financial institution. Having this information should allow to differentiate between bank types in terms of punishment, and to reduce the noisy adverse selection problem, so that especially more prudent banks can be induced to report less conservatively, and to better exploit their regulatory capital. In particular, when the bank's short-term risk attitude θ_t is correlated over time then the above discussion suggests that especially banks with low capital ratios should be induced to report less conservatively.

7. Summary and policy implications

Recent empirical studies suggest a *pronounced conservatism in the risk figures* disclosed by financial institutions vis-a-vis regulatory authorities. In a simple model involving informational frictions between financial institution and regulator, we pointed out that this observation can be explained as the *consequence of an adverse selection problem*, whose mechanics induce less risk-prone financial institutions to use overly conservative and less profitable investment strategies, and to report nevertheless maximum risks.

The recent evidence gains utmost relevance in view of the recent Basel II proposals that leave market risk supervision essentially in its 1996 design. The analysis suggests that *supervisors should want to disencourage overly conservative risk reporting.* This could be achieved easily by implementing statistical tests that measure the model's conservatism (*two-sided backtesting*). Moreover, it appears that an appropriate incentive system should be linked to the historical capital ratio of the respective bank or securities firm. Specifically, non-conservatism should be rewarded especially when the bank has a history of low capital exploitation.

Introducing these additional incentives should have a number of desirable consequences. The first and obvious is that *risk figures may be more accurate*, and give a more transparent perspective on the stability of the financial sector. More importantly, however, is the effect that prudent banks would have higher marginal costs from conservative risk reporting, and therefore better incentives to exploit their regulatory risk capital. This would *mitigate the adverse selection problem* and thereby improve the overall efficiency of the regulatory mechanism.

References

Basel Committee on Banking Supervision (1996a): "Overview of the Amendment to the Capital Accord to Incorporate Market Risk," Bank for International Settlements.

Basel Committee on Banking Supervision (1996b): "Amendment to the Capital Accord to Incorporate Market Risk," Bank for International Settlements.

Basel Committee on Banking Supervision (1996c): "Supervisory Framework for the Use of 'Backtesting' in Conjunction with the Internal Models Approach to Market Risk Capital Requirements," Bank for International Settlements.

Basel Committee on Banking Supervision (1999): "Performance of Models-Based Capital Charges for Market Risk," Bank for International Settlements.

Basel Committee on Banking Supervision (2001): "The Internal Ratings-Based Approach, Consultative Document" Bank for International Settlements.

Baumol, W. J. (1963): "An Expected Gain-Confidence Limit Criterion For Portfolio Selection," *Management Science* **10 (1)**, 174-182.

Berger, A., R. Herring, and G. Szegö (1995): "The Role of Capital in Financial Institutions," *Journal of Banking and Finance* **19**, 393-430.

Berkowitz, J., and J. O'Brien (2001): "How Accurate are Value-at-Risk Models at Commercial Banks?" *Discussion Paper*, Federal Reserve Board.

Bhattacharya, S., Boot, A., and A. Thakor (1998): "The Economics of Bank Regulation," *Journal of Money, Credit, and Banking* **30 (4)**, 745-770.

Bhattacharya, S., and A. Thakor (1993): "Contemporary Banking Theory," *Journal of Financial Intermediation* **3**, 2-50.

Blum, J. (1999): "Do Capital Adequacy Requirements Reduce Risks in Banking?" *Journal of Banking and Finance* **23**, 755-771.

Blum, J., and M. Hellwig (1995): "The Macroeconomic Implications of Capital Adequacy Requirements for Banks," *European Economic Review* **39**, 739-49.

Caillaud, B., R. Guesnerie, and P. Rey (1992): "Noisy Observation in Adverse Selection Models," *Review of Economic Studies* **59**, 595-615.

Ewerhart, C. (2000): "Market Risks, Internal Models and Optimal Regulation: Does Backtesting Induce Banks to Report Their True Risks," *SFB 504 Discussion Paper 00-22*.

Freixas, X., and J.-C. Rochet (1997): *Microeconomics of Banking*, MIT Press, Cambridge, Massachusetts.

Gizycki, M., and N. Hereford (1998): "Assessing the Dispersion in Banks' Estimates of Market Risk: The Results of a Value-at-Risk Survey," *Discussion Paper* 1, Australian Prudential Regulation Authority.

Hart, O., and D. Jaffee (1974): "On the Application of Portfolio Theory of Depository Financial Intermediaries," *Review of Economic Studies* **41**, 129-147.

Hellwig, M., and M. Staub (1996): "Capital Requirements for Market Risks Based on Inhouse Models – Aspects of Quality Assessment," *Swiss Journal of Economics and Statistics* **132 (4/2)**, 755-76.

Jorion, P. (1995): *Value at Risk*, McGraw-Hill, New York.

Keeley, M. C., and Furlong, F. T. (1990): “A Re-Examination of the Mean-Variance Analysis of Bank Capital Regulation,” *Journal of Banking and Finance* **14**, 69-84.

Kim, D., and Santomero, A. M. (1988): “Risk in Banking and Capital Regulation,” *Journal of Finance* **43 (5)**, 1219-33.

Koehn, M., and Santomero, A. M. (1980): “Regulating of Bank Capital and Portfolio Risk,” *Journal of Finance* **35 (5)**, 1235-44.

Lucas, A. (2001): “Evaluating the Basel Guidelines for Backtesting Banks’ Internal Risk Management Models, *Journal of Money, Credit, and Banking* **33 (3)**, 826-841.

Milgrom, P., and J. Roberts (1990): Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities, *Econometrica* **58 (6)**, 1255-1277.

Pyle, D. (1971): “On the Theory of Financial Intermediation,” *Journal of Finance* **26**, 737-747.

Rime, B. (2001): Capital Requirements and Bank Behavior,” *Journal of Banking and Finance* **25 (4)**, 789-805.

Rochet, J.-C. (1992): “Capital Requirements and the Behaviour of Commercial Banks,” *European Economic Review* **36**, 1137-1178.

Number of exceptions k	Capital multiplier l_1
4 or less	3.00
5	3.40
6	3.50
7	3.65
8	3.75
9	3.85
10 or more	4.00

Table 1. The Basle Committee incentive scheme

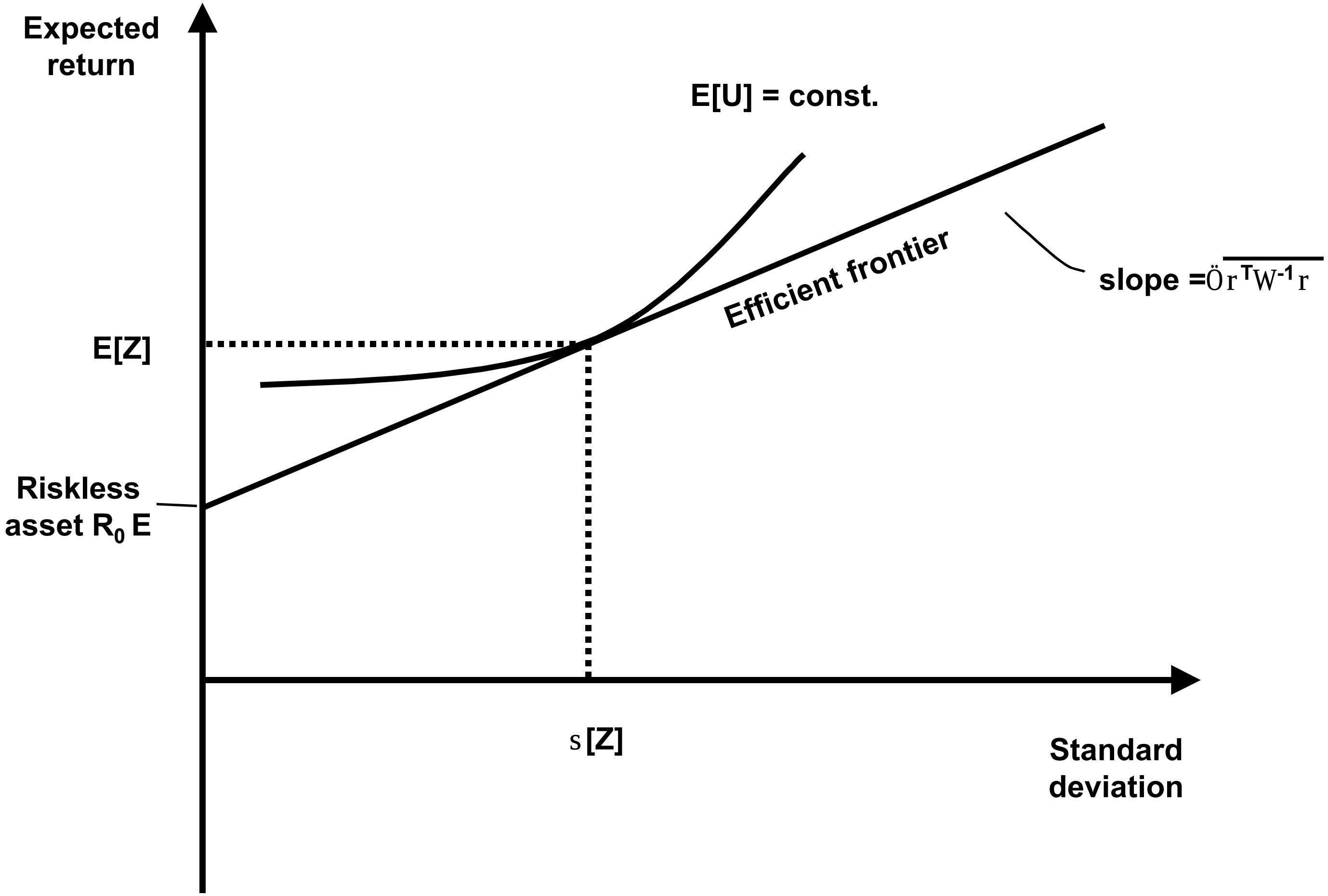

Figure 1: CAPM investment opportunities

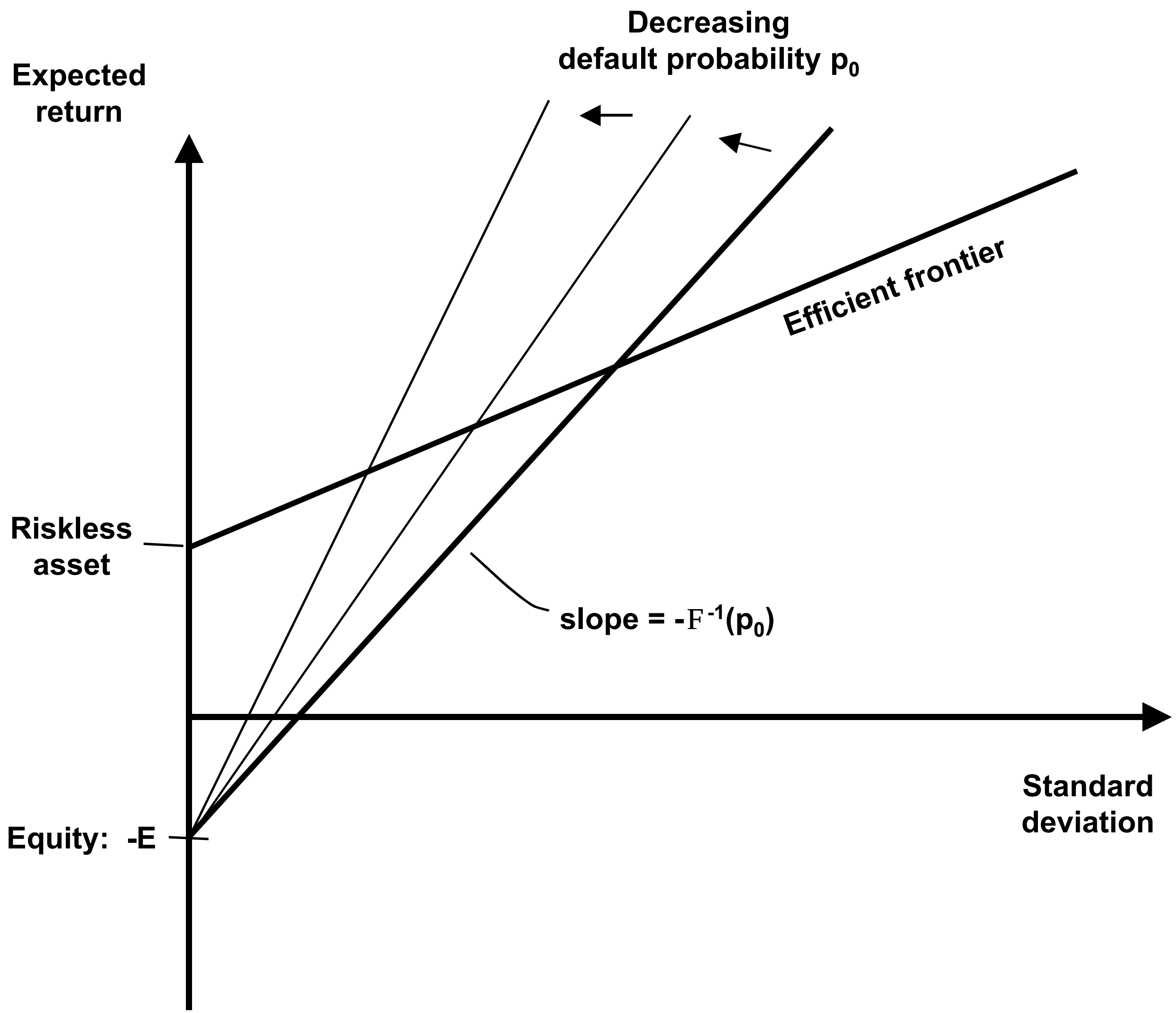

Figure 2: Iso-default-probability curves

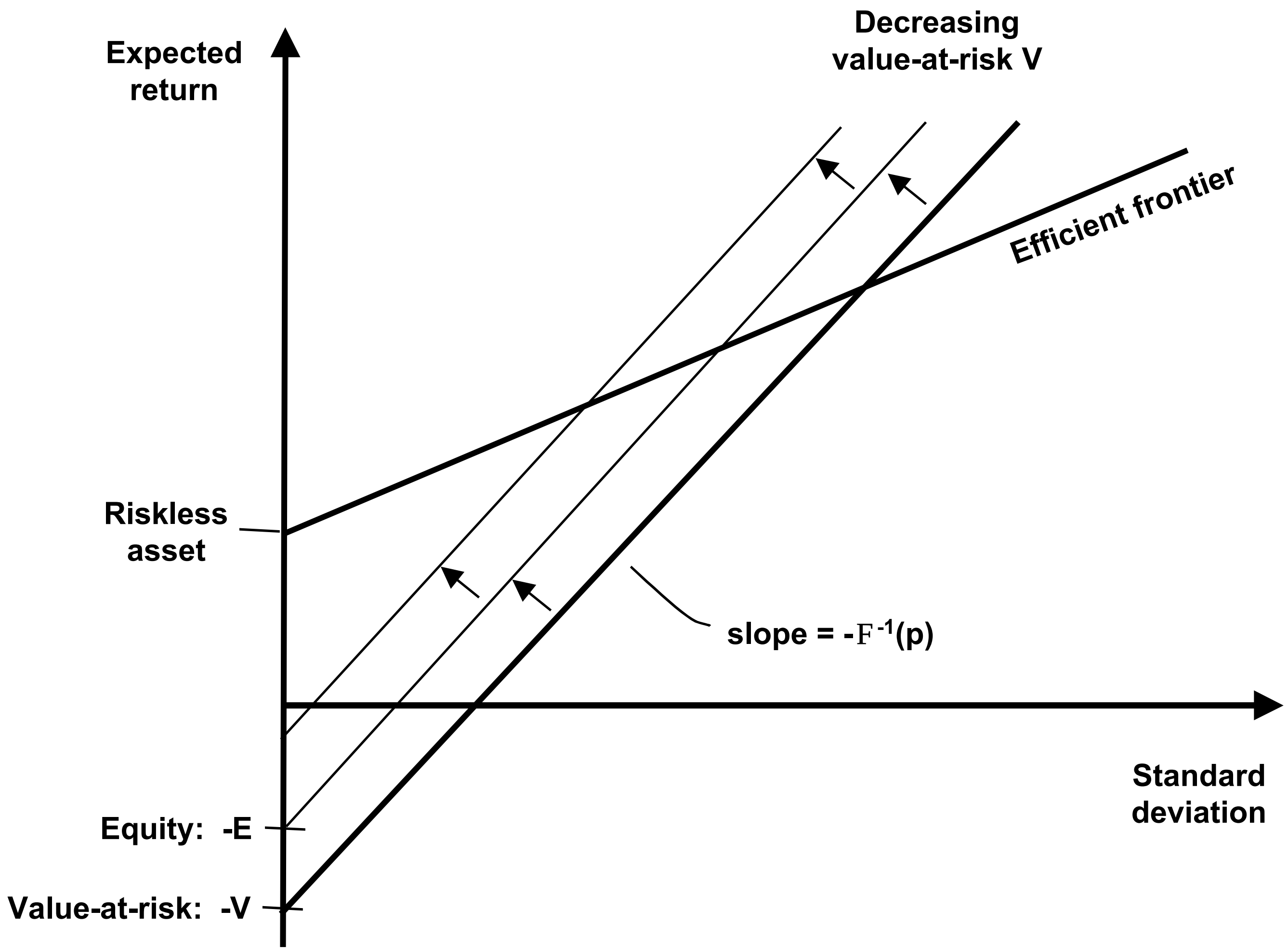

Figure 3: Iso-value-at-risk curves

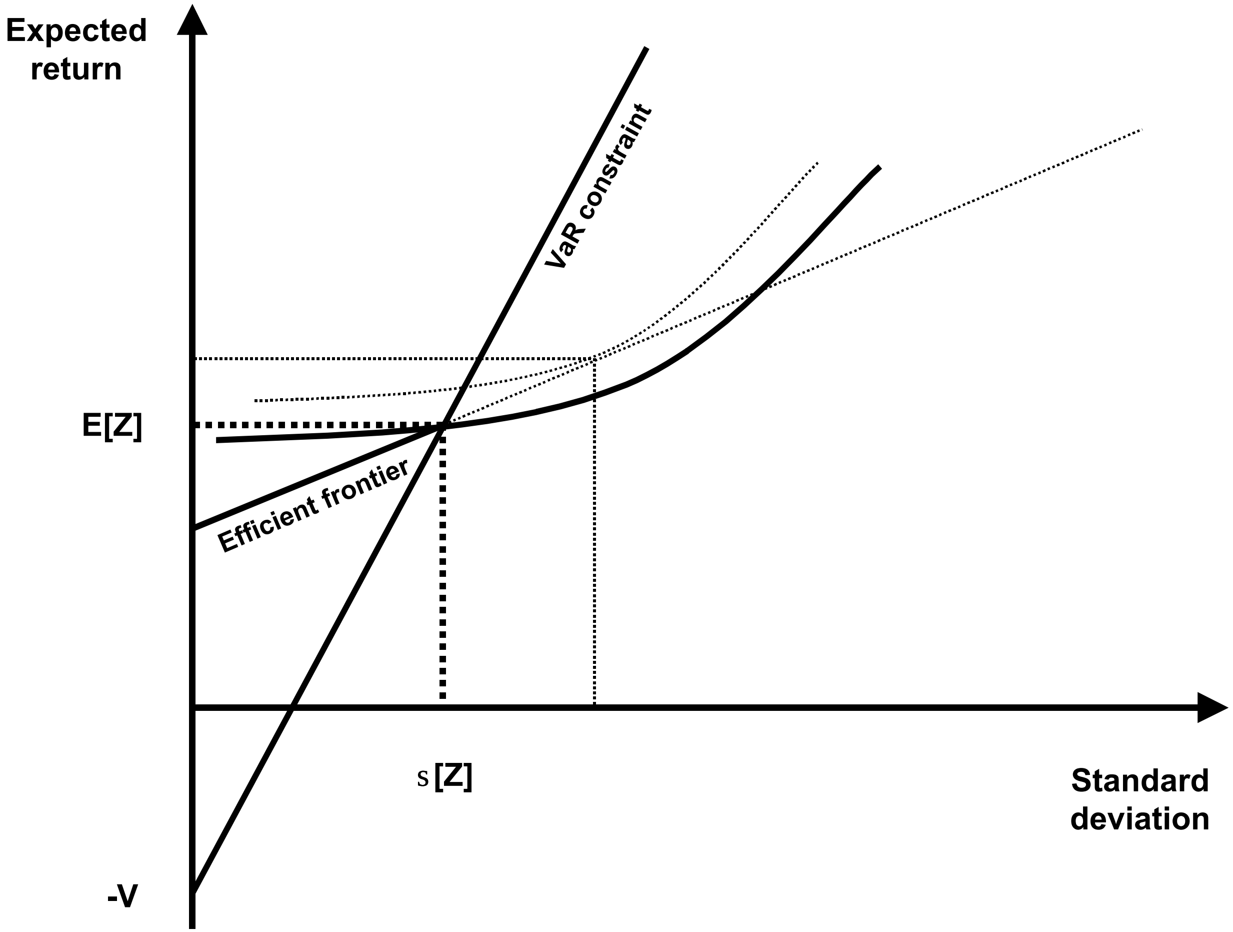

Figure 4: Value-at-risk regulation

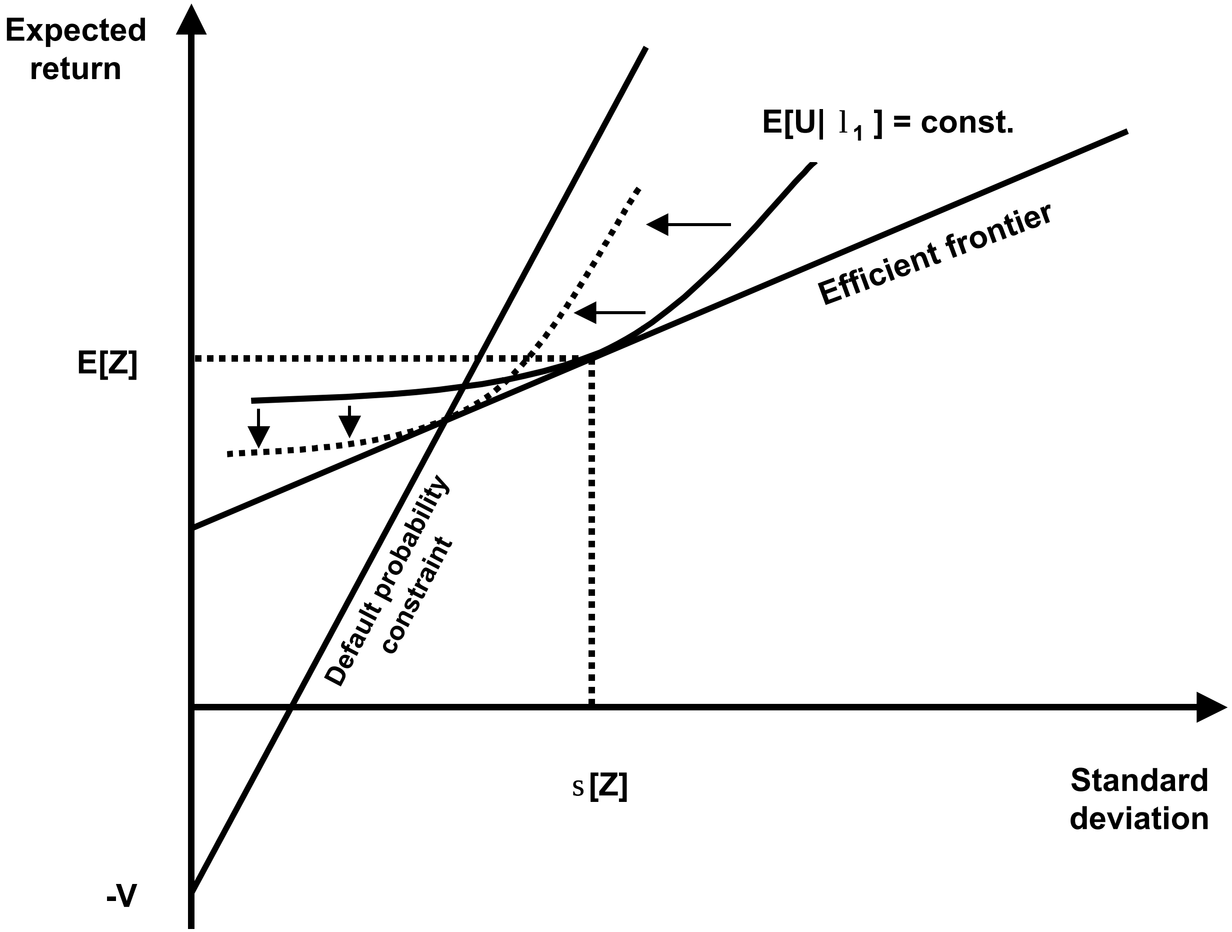

Figure 5: Asymmetric information: need for incentives