Is regulating the solvency of banks counter-productive?*

by
Peter Zweifel, Dieter Pfaff, and Jochen Kühn

This version: November 2008

Peter Zweifel, Socioeconomic Institute, University of Zurich (Switzerland), Hottingerstrasse 10, CH-8032 Zürich, pzweifel@soi.uzh.ch (corresponding author)
Dieter Pfaff, Institute for Accounting and Controlling, University of Zurich (Switzerland), Plattenstr. 14, CH–8032 Zürich, dieter.pfaff@irc.uzh.ch
Jochen Kühn, Volkenbachstr. 4, D-79798 Jestetten (Germany), mail@jochen-kuehn.de

Abstract

This paper contains a critique of solvency regulation such as imposed on banks by *Basel I* and *II*. It argues that banks seeking to maximize rate of return on risk-adjusted capital (RORAC) aim at an optimal level of solvency because on the one hand, solvency S lowers the cost of re-financing; on the other, it ties costly capital. In period 1, exogenous changes in mean returns $d\bar{\mu}$ and in volatility occur, causing optimal adjustments $dS^*/d\bar{\mu}$ and $dS^*/d\bar{\sigma}$ in period 2. Since banks reallocate their assets with certain μ and σ values in response to the changed solvency level, an endogenous trade-off with slope $d\mu/d\sigma$ results in period 3. Both *Basel I* and *II* are shown to modify this slope, inducing at least some banks to opt for a higher value of σ in certain situations. Therefore, this type of solvency regulation can prove counter-productive.

JEL codes: G15, G21, G28, L51

Key words: regulation, banks, solvency, *Basel I*, *Basel II*

* The authors acknowledge helpful suggestions and criticisms by Sebastian van Dahlen and Thomas Müller (Bank for International Settlements), Christian Ewerhart (University of Zurich), and Huan Wang (University of Zurich).

1 Introduction

The concept of risk-adjusted capital was originally developed for the performance measurement of a bank's trading division. However, in the meantime return on risk-adjusted capital (RORAC) has increasingly become the benchmark for assessing an entire bank's performance and governance. At the same time, public regulators are concerned about solvency to ensure the continuity of a bank's operations. This paper deals with the conflict between optimization of RORAC (which also implies an optimal solvency level) and exogenously imposed solvency levels, taking *Basel I* and *Basel II* as the example. It does so by distinguishing exogenous $\left(\bar{\mu},\bar{\sigma}\right)$ and endogenous (μ,σ) components in expected returns and volatility of returns. In a first period, exogenous shocks ($d\bar{\mu}, d\bar{\sigma}$)impinge on the bank's optimum. A typical cause could be investments made in the previous period that turn out to have a lower rate of return or a higher volatility than expected. In the second period, banks optimally adjust their solvency levels by $dS^*/d\bar{\mu}$ and $dS^*/d\bar{\sigma}$, respectively. In the third period, they target new values of μ and σ on a perceived efficiency frontier with slope $d\mu/d\sigma$, which is implied by $dS^*/d\bar{\mu}$, $dS^*/d\bar{\sigma}$, and the fact that a single net adjustment dS^* occurred.

This efficiency frontier is modified by solvency regulation such as *Basel I* and *II*. It will be shown that *Basel I* simply sets a lower limit on (partially risk-adjusted) capital, neglecting that this limit should reflect both $\bar{\mu}$ and $\bar{\sigma}$ and the fact that the relationship between solvency and capital depends on $\bar{\mu}$ and $\bar{\sigma}$ as well. As to *Basel II*, it addresses solvency directly but fails to take into account the fact that a bank that initially just met this standard must come up with additional capital when $\bar{\mu}$ falls or $\bar{\sigma}$ increases. Therefore both *Basel I* and *II* modify the efficiency frontier $d\mu/d\sigma$ as perceived by regulated banks. While one might expect that these regulations cause the slope of the frontier to become steeper (thus inducing banks to opt for lower μ and lower σ), it turns out that the opposite can be the case. Indeed, both, *Basel I* and *II* may have the unexpected (and presumably counter-productive) consequence of causing at least some banks in some constellations to opt for a higher value of σ (i.e. higher volatility of the rate of return on their assets) than without it.

This paper is structured as follows. Section 2 contains a review of the pertinent literature to conclude that solvency regulation indeed may serve to avoid negative externalities. In section 3, a higher level of solvency is found to have two effects for banks aiming to maximize

RORAC. On the one hand, it serves to lower their cost of refinancing; on the other, it ties capital that would have other, more productive uses. In section 4, comparative statics are used to derive optimal second-period adjustments of solvency to exogenous first-period changes in rate of return $dS^*/d\bar{\mu}$ and in volatility $dS^*/d\bar{\sigma}$, respectively. The net adjustment results in the bank's third-period endogenous tradeoff $d\mu/d\sigma$. Next, the regulations imposed by *Basel I* and *II* are introduced as parameter restrictions to show how $d\mu/d\sigma$ is modified, possibly resulting in regulated banks opting for a higher value of σ than on their own. A summary and conclusions follow in section 5.

2 Literature review

The solvency regulation of banks has traditionally been justified by the external costs of insolvency, especially in the guise of a bank run (Diamond and Dybvig, 1983). This view was challenged by the proponents of the Capital Asset Pricing Model, who emphasized that for well-diversified investors, the solvency of a bank does not constitute a reasonable objective. They are concerned with expected profitability, possibly adjusted for the degree to which the bank's profitability systematically varies with the capital market (the Beta of the Capital Asset Pricing Model). By way of contrast, for little-diversified investors (among them, ordinary consumers holding deposits with the bank), the bank's overall risk is relevant, which importantly includes the risk of insolvency [Goldberg and Hudgins (1996), Park and Peristiani (1998), Jordan (2000), Goldberg and Hudgins (2002)]. Option Pricing Theory shows that due to their limited liability, shareholders of the bank in fact have a put option that is written by the other stakeholders (notably creditors) of the bank [Merton (1974), Jensen and Meckling (1976), Merton (1977)].

When a solvency risk materializes, internal and external costs need to be distinguished. Internal costs are borne by the bank's shareholders, who see the value of their shares drop to zero unless the bank is in business again. However, in view of the loss of reputation, this re-entry would meet with high barriers to entry [Smith and Stulz (1985), pp. 395-396, Stulz (1996), pp. 9-12]. In addition, insolvency has external costs (i.e. costs not borne by the insolvent bank). First, the insolvency may trigger a bank run [Diamond and Dybvig (1983), Jacklin and Bhattacharya (1988), Bauer and Ryser (2004)]. Depositors who are late to withdraw their funds stand to lose part of their assets. Some of these depositors may be banks themselves;

therefore, the insolvent bank may drive other financial institutions into bankruptcy, causing substantial external costs [Lang and Stulz (1992), Furfine (2003)]. Second, investors in the capital market at large often are affected as well. A bank that becomes insolvent causes owners and creditors of banks in general to re-evaluate the estimated risk of insolvency. In response to the revised estimate, they demand a higher rate of interest from their banks, driving up the cost of refinancing. There is a substantial body of empirical research substantiating this claim [Flannery and Sorescu (1996), Park and Peristiani (1998), Covitz et al. (2004)].

Conclusion 1: A solvency level that is deemed optimal by the individual bank is too low from a societal perspective because an insolvency causes substantial external costs.

It may be worthwhile to emphasize that this conclusion does not suffice to justify public regulation to ensure solvency. One would have to first examine whether the expected benefit of the intervention exceed its expected cost. An important component of this cost is caused by behavioral adjustments that are not intended. The present contribution belongs to this tradition of research, which dates back at least to Koehn and Santomero (1980). Characterizing a bank by its utility function and assuming it to optimize a portfolio containing both assets and liabilities, they find that imposing a simple equity-to-assets ratio constraint is ineffective on average. Relatively safe banks become safer, while risky ones increase their risk position to make up for decreased leverage. In Kim and Santomero (1988), emphasis is on the choice of appropriate risk weights in the determination of what has become since 'Risk-Adjusted Capital'. Here, the cost of regulation derives from non-optimal risk weights.

In Rochet (1992), banks choose their asset portfolio taking into account limited liability, which may cause them to become risk-lovers. This makes imposing minimum capital requirements necessary to prevent them from choosing very inefficient portfolios. However, the effectiveness of this regulation is not guaranteed at all. John, Saunders, and Senbet (2000) show that U.S. capital-based regulation introduced in 1991 may fail to prevent bank managers from shifting risk to outside financiers unless features of their compensation plans are taken into account along with the opportunity set of asset investments. More recently, Repullo (2004) explicitly has dealt with *Basel II* in the context of an imperfectly competitive market. He derives conditions for two Nash equilibria to obtain, one in which banks invest in riskless and another where they invest risky assets. While capital requirements on risky assets do enlarge the parameter space of the 'prudent' equilibrium, depositors bear the burden of regulation in the guise of lower interest rates. That is also the reason why in Repullo (2004) capital requirements are in general effective in preventing excessive risk-taking by banks. Furthermore, it is shown that *Basel II* permits a reduction in the overall amount of capital required by regulation compared to *Basel I*.

The present contribution differs from the earlier literature in two ways. First, it clearly distinguishes between the earlier *Basel I* and the more refined *Basel II* regulation, showing that the more recent variant may have unintended consequences only for a subset of banks rather than all of them. In this respect, this work elaborates on and refines the contributions by Kim and Santomero (1988) as well as Rochet (1992). The second distinguishing feature of this paper is its emphasis on dynamics in the following way. Whereas earlier contributions analyzed optima or [in the case of Repullo (2004)] equilibria, here banks adjustment to shocks from a previous period take center stage. It is this adjustment that will be shown to be conditioned by solvency regulation of the *Basel I* and *II* type. In return, welfare implications will not be spelled out; rather, the fact that banks may be induced to act against the stated intentions of the regulator will be highlighted.

3 Optimal solvency in a model of bank behavior

Let the bank's management maximize the (expected) rate of return on risk-adjusted capital (RORAC). For simplification, the model is couched in terms of nonstochastic variables; still, risk enters through the fact that a higher level of solvency S enables the bank to obtain funds at a lower rate of interest paid on deposits r_D. Therefore, one has

$$r_D = r_D(S), \quad \text{with } \frac{\partial}{\partial S} r_D(S) < 0 \text{ and } \frac{\partial^2}{\partial S^2} r_D(S) > 0. \qquad (1)$$

The business portfolio (of fixed size for simplicity) has an expected return $\hat{\mu}$ and volatility $\hat{\sigma}$, comprising an exogenous $(\overline{\mu}, \overline{\sigma})$ and an endogenous (μ, σ) component such that $\hat{\mu} = \overline{\mu} + \mu$ and $\hat{\sigma} = \overline{\sigma} + \sigma$. These parameters determine the level of risk capital C required to attain a certain solvency level,

$$C = C\left(\hat{\mu}, \hat{\sigma}, S\right) > 0. \qquad (2)$$

Clearly, an exogenous increase in the rate of return serves to decrease this capital requirement, while an exogenous increase in volatility or targeted solvency serves to increase it,

$$\frac{\partial}{\partial\bar{\mu}}C(\cdot)<0, \frac{\partial}{\partial\bar{\sigma}}C(\cdot)>0, \frac{\partial}{\partial S}C(\cdot)>0. \tag{3}$$

It is further assumed that while risk capital requirements increase with derived solvency, the increase is mitigated when the exogenous component of the expected rate of return $\bar{\mu}$ on capital markets is high, making it easier for the bank to achieve the solvency margin. By way of contrast, a higher solvency level calls for even more risk capital C when volatility of investment returns $\bar{\sigma}$ exogenously increases. Therefore, one has

$$\frac{\partial^2}{\partial\bar{\mu}\partial S}C(\cdot)<0 \text{ and } \frac{\partial^2}{\partial\bar{\sigma}\partial S}C(\cdot)>0. \tag{4}$$

Risk capital is invested at a risk-free interest rate r_f. Operating costs, taxes and restrictions such as equity capital regulations are being disregarded. On the basis of these assumptions, RORAC can be expressed as follows,

$$\text{RORAC} = \frac{\left(\hat{\mu}-r_D(S)\right)D+r_f C\left(\hat{\mu},\hat{\sigma},S\right)}{C\left(\hat{\mu},\hat{\sigma},S\right)}. \tag{5}$$

Since the volume of the business portfolio is assumed to be constant, the maximization of RORAC leads to the following first-order condition for optimal solvency,

$$-\frac{\partial}{\partial S}r_D[S^*]=\frac{\hat{\mu}-r_D[S^*]}{C[\cdot,S^*]}\frac{\partial}{\partial S}C[\cdot,S^*], \tag{6}$$

with the bracket notation pointing to the fact that the endogenous determinant S has to be evaluated at its optimal level. Equation (6) can be interpreted as follows. It is optimal for a bank to weigh the increased solvency's favorable marginal effect on the cost of refinancing (left-hand side of the equation, called marginal return of solvency in terms of risk cost) against its marginal downside effect on solvency (right-hand side, called the marginal cost of solvency). The marginal cost of solvency consists of two interacting components. First, solvency ties costly capital C. Secondly however, this cost is particularly high when the rate of return achievable $\hat{\mu} = \overline{\mu} + \mu$ exceeds by far the bank's refinancing cost r_D.

Conclusion 2: A bank that seeks to maximize its rate of return on risk-adjusted capital is predicted to optimize its solvency level by balancing its marginal return in terms of reduced cost of refinancing cost against its marginal cost in terms of tied capital and foregone returns.

Equation (6) makes clear that the optimal adjustment to an exogenous change in solvency requirements is not given once and for all but importantly depends on the risk-return profile inherited by the banks from the past.

Before substantiating this claim, it is worthwhile to note that fixing a certain solvency level to be adhered to at all times does not only entail disadvantages. One advantage is simplicity, although the management of a bank may be hard put to operationalize 'level of solvency' in all circumstances. Second, the conventional policy in fact makes the cost of (re) financing independent of investment decisions, permitting separation of the bank's lending and borrowing policies, which again results in an important simplification of management tasks. On the downside, uniform regulation creates a similarity in the decision-making situation of regulated firms, which usually results in a type of implicit collusion limiting competition.

4 Determination of the perceived efficiency frontier

The purpose of this section is to show how conventional solvency regulation restricts bank behavior. Specifically, capital requirements will be seen to affect a bank's optimal risk-return trade-off. In the first period, expected returns and volatility change exogenously by $d\overline{\mu}$ and $d\overline{\sigma}$, respectively. During the second period, solvency adjusts optimally to these stocks according to $dS^*/d\overline{\mu}$ and $dS^*/d\overline{\sigma}$. In the third period, net adjustment dS^* from the second period causes banks to choose μ and σ anew through adjusting their assets and liabilities. These adjustments result in a perceived endogenous efficiency frontier with slope $d\mu/d\sigma$ that is relevant for allocative decisions in the third period.

4.1 Perceived efficiency frontier without regulation

The starting point is the first-order condition (6) for a maximum, rewritten

$$\frac{dR}{dS} = 0\,, \tag{7}$$

with R shorthand for RORAC. Now consider a shock $d\bar{\mu}$ disturbing this first-order condition. Since (7) is satisfied at the new optimum as well, optimal adjustment by dS^* must be such as to neutralize the shock, implying

$$\frac{\partial^2 R}{\partial S^2} dS^* + \frac{\partial^2 R}{\partial S \partial \bar{\mu}} d\bar{\mu} = 0. \tag{8}$$

This can be solved to obtain,

$$\frac{dS^*}{d\bar{\mu}} = -\frac{\partial^2 R / \partial S \partial \bar{\mu}}{\partial^2 R / \partial S^2}\,. \tag{9}$$

Since $\partial^2 R / \partial S^2 < 0$ in the neighborhood of a maximum, sgn $\left[\partial^2 R / \partial S \partial \bar{\mu}\right]$ determines sgn $\left[dS^* / d\bar{\mu}\right]$. Differentiating (6) w.r.t. $\bar{\mu}$, one has

$$\operatorname{sgn}\left[\frac{\partial^2 R}{\partial S \partial \bar{\mu}}\right] = \operatorname{sgn}\left[\frac{\partial^2}{\partial S \partial \bar{\mu}}\left[r_D\right] + \frac{(1-\partial r_D / \partial \bar{\mu})C - \left(\hat{\mu} - r_D\right)\partial C / \partial \bar{\mu}}{C^2} \cdot \frac{\partial C}{\partial S} + \frac{\hat{\mu} - r_D}{C} \cdot \frac{\partial^2 C}{\partial S \partial \bar{\mu}}\right]. \tag{10}$$

Using (6) to substitute for $\left[\left(\hat{\mu} - r_D\right)/C\right] \cdot \partial C / \partial S$ and $\left(\hat{\mu} - r_D\right)/C$, respectively, and multiplying by $\bar{\mu} \cdot S / r_D$, one obtains

$$\begin{aligned}\operatorname{sgn}\left[\frac{\partial^2 R}{\partial S \partial \bar{\mu}}\right] = \operatorname{sgn}\Bigg[&\frac{\partial}{\partial \bar{\mu}}\left[\frac{\partial r_D}{\partial S} \cdot \frac{S}{r_D}\right]\bar{\mu} + \frac{S}{C} \cdot \frac{\partial C}{\partial S} \cdot \frac{\bar{\mu}}{r_D}\left(1 - \frac{\partial r_D}{\partial \bar{\mu}}\right) + \frac{\partial r_D}{\partial S} \cdot \frac{S}{r_D} \cdot \frac{\partial C}{\partial \bar{\mu}} \cdot \frac{\bar{\mu}}{C} \\ &- \frac{\partial r_D}{\partial S} \cdot \frac{S}{r_D} \cdot \frac{\bar{\mu}}{\partial C / \partial S} \cdot \frac{\partial}{\partial \bar{\mu}}\left[\frac{\partial C}{\partial S}\right]\Bigg].\end{aligned} \tag{11}$$

This can be rewritten to become

$$\begin{aligned}&\operatorname{sgn}\left[\frac{\partial^2 R}{\partial S \partial \bar{\mu}}\right] = \\ &\operatorname{sgn}\left[\frac{\partial}{\partial \bar{\mu}}\left[\frac{\partial r_D}{\partial S} \cdot \frac{S}{r_D}\right]\bar{\mu} + \left(\frac{\partial C}{\partial S} \cdot \frac{S}{C}\right)\left(1 - \frac{\partial r_D}{\partial \bar{\mu}}\right)\frac{\bar{\mu}}{r_D}\right] + \left(\frac{\partial r_D}{\partial S} \cdot \frac{S}{r_D}\right)\left[\left(\frac{\partial C}{\partial \bar{\mu}} \cdot \frac{\bar{\mu}}{C}\right) - \frac{\bar{\mu}}{\frac{\partial C}{\partial S} \cdot \frac{S}{C}} \cdot \frac{\partial}{\partial \bar{\mu}}\left[\frac{\partial C}{\partial S} \cdot \frac{S}{C}\right]\right]\Bigg].\end{aligned} \tag{12}$$

Using elasticity notation, this becomes

$$\operatorname{sgn}\left[\frac{dS^*}{d\bar{\mu}}\right]=\operatorname{sgn}\left[\frac{\partial^2 R}{\partial S\partial\bar{\mu}}\right]=\operatorname{sgn}\left[\underset{(+)}{\frac{\partial}{\partial\bar{\mu}}\left[e\left(r_D,S\right)\right]}\bar{\mu}+\underset{(+)}{e\left(C,S\right)}\left(\frac{\bar{\mu}}{r_D}-\underset{(-)}{e(r_D,\bar{\mu})}\right)\right.$$

$$\left.+\underset{(-)}{e\left(r_D,S\right)}\left(\underset{(-)}{e\left(C,\bar{\mu}\right)}-\frac{\bar{\mu}}{\underset{(+)}{e\left(C,S\right)}}\cdot\underset{(-)}{\frac{\partial}{\partial\bar{\mu}}\left[e\left(C,S\right)\right]}\right)\right],$$

(13)

with $e\left(r_D,S\right):=\frac{\partial r_D}{\partial S}\cdot\frac{S}{r_D}<0,\ e\left(C,S\right):=\frac{\partial C}{\partial S}\cdot\frac{S}{C}>0,\ e(C,\bar{\mu}):=\frac{\partial C}{\partial\bar{\mu}}\cdot\frac{\bar{\mu}}{C}<0$, and $e\left(r_D,\bar{\mu}\right):=\frac{\partial r_D}{\partial\bar{\mu}}\cdot\frac{\bar{\mu}}{r_D}<0$.

These elasticities are treated as constant except with respect to $\bar{\mu}$ and $\bar{\sigma}$. As to signs, the elasticity $e(r_D, S)$ is certainly negative. Next, $e(C, S) > 0$ because efforts to increase solvency typically call for additional reserves or capital. Increased profitability $\bar{\mu}$ permits the bank to reduce capital requirements, thus $e\left(C,\bar{\mu}\right)<0$; it also helps to obtain outside financing at a lower interest rate, implying $e\left(r_D,\bar{\mu}\right)<0$. The signs of the two derivatives $\partial/\partial\bar{\mu}\left[e\left(r_D,S\right)\right]>0$ and $\partial/\partial\bar{\mu}\left[e\left(C,S\right)<0\right]$ can be justified as follows. While a higher expected rate of return $\bar{\mu}$ makes the bank become a better risk, permitting it to obtain outside finance at more favorable conditions, this effect will taper out sooner or later, such that $\partial/\partial\bar{\mu}\left[e\left(r_D,S\right)\right]>0$. As to $e(C, S) > 0$, an increase of $\bar{\mu}$ is likely to weaken the connection between solvability and capital requirements because the regulator perceives insolvency risk to be reduced, hence $\partial/\partial\bar{\mu}\left[e\left(C,S\right)\right]<0$. Using these signs in equation (13), one obtains

$$\frac{dS^*}{d\bar{\mu}}>0 \text{ for } \bar{\mu}\to 0;\ \frac{dS^*}{d\bar{\mu}}\lesseqgtr 0 \text{ for } \bar{\mu}\to\infty. \tag{14}$$

In full analogy,

$$\operatorname{sgn}\left[\frac{\partial^2 R}{\partial S\partial\bar{\sigma}}\right]=\operatorname{sgn}\left[\frac{\partial^2}{\partial S\partial\bar{\sigma}}\left[r_D\right]-\left(C\cdot\frac{\partial r_D}{\partial\bar{\sigma}}+(\bar{\mu}-r_D)\cdot\frac{\partial C}{\partial\bar{\sigma}}\right)\cdot\frac{1}{C^2}\cdot\frac{\partial C}{\partial S}+\frac{\bar{\mu}-r_D}{C}\cdot\frac{\partial^2 C}{\partial S\partial\bar{\sigma}}\right]. \tag{15}$$

Using (6) again, one obtains after multiplying through by $\sigma\cdot S/r_D$,

$$\operatorname{sgn}\left[\frac{\partial^2 R}{\partial S \partial \bar{\sigma}}\right] = \operatorname{sgn}\left[\frac{\partial}{\partial \bar{\sigma}}\left[\frac{\partial r_D}{\partial S}\cdot\frac{S}{r_D}\right]\bar{\sigma} - \left(\frac{\partial C}{\partial S}\cdot\frac{S}{C}\right)\cdot\left(\frac{\partial r_D}{\partial \bar{\sigma}}\cdot\frac{\bar{\sigma}}{r_D}\right)\right.$$
$$\left. + \left(\frac{\partial r_D}{\partial S}\cdot\frac{S}{r_D}\right)\cdot\left(\frac{\partial C}{\partial \bar{\sigma}}\cdot\frac{\bar{\sigma}}{C} - \frac{\bar{\sigma}}{\frac{\partial C}{\partial S}\cdot\frac{S}{C}}\cdot\frac{\partial}{\partial \bar{\sigma}}\left[\frac{\partial C}{\partial S}\cdot\frac{S}{C}\right]\right)\right].$$

(16)

Using elasticity notation once more, this becomes

$$\operatorname{sgn}\left[\frac{dS^*}{d\bar{\sigma}}\right]\operatorname{sgn}\left[\frac{\partial^2 R}{\partial S \partial \bar{\sigma}}\right] = \operatorname{sgn}\left[\underset{(+)}{\frac{\partial}{\partial \bar{\sigma}}}\left[e\left(\underset{(+)}{r_D, S}\right)\right]\bar{\sigma} + \underset{(+)}{e(C,S)}\cdot\left(-\underset{(+)}{e(r_D,\bar{\sigma})}\right)\right.$$
$$\left. + \underset{(-)}{e(r_D,S)}\cdot\left(\underset{(+)}{e(C,\bar{\sigma})} - \frac{\bar{\sigma}}{\underset{(+)}{e(C,S)}}\cdot\underset{(+)}{\frac{\partial}{\partial \bar{\sigma}}}\left[e(C,S)\right]\right)\right]. \tag{17}$$

With increasing volatility, the interest rate to be paid for outside financing r_D increases because creditors bear more risk while not participating in profits $\left[e(r_D,\bar{\sigma}) > 0\right]$. With increasing volatility, the regulator is likely to require more capital, hence $e(C,\bar{\sigma}) > 0$. The signs of the additional derivatives can be justified as follows. While $e(r_D, S) < 0$, effectiveness of solvency effort diminishes when the bank is judged more risky due to an increase in $\bar{\sigma}$, implying $\partial / \partial\bar{\sigma}\left[e(r_D, S)\right] > 0$. Similarly, the regulator is likely to require that solvency efforts be increasingly backed up by additional reserves, thus $\partial / \partial\bar{\sigma}\left[e(C, S)\right] > 0$. Using these signs in equation (17), one obtains

$$\frac{dS^*}{d\bar{\sigma}} < 0 \text{ if } \bar{\sigma} \to 0; \quad \frac{dS^*}{d\bar{\sigma}} > 0 \text{ for } \bar{\sigma} \to \infty . \tag{18}$$

Equations (14) and (18) define the predicted second-period adjustments in response to shocks that occurred in the first period. However, there can be only one net value dS^* that in turn triggers adjustments in μ and σ during the third period. By reallocating its assets and liabilities, the bank moves along a perceived efficiency frontier with slope $d\mu/d\sigma$. This slope can be obtained by dividing (17) by (13), yielding

$$\left.\frac{d\mu}{d\sigma}\right|_{S^*} = \frac{\dfrac{\partial}{\partial\bar{\sigma}}\left[\overset{(+)}{e\left(r_D,S\right)}\right]\bar{\sigma} - \overset{(+)}{e\left(C,S\right)}\cdot \overset{(+)}{e\left(r_D,\bar{\sigma}\right)} + \overset{(-)}{e\left(r_D,S\right)}\cdot \overset{(+)}{e\left(C,\bar{\sigma}\right)} - \overset{(-)}{e\left(r_D,S\right)}\cdot \dfrac{\overset{(+)}{\bar{\sigma}}}{e\left(C,S\right)}\cdot \dfrac{\overset{(+)}{\partial}}{\partial\bar{\sigma}}\left[e\left(C,S\right)\right]}{\dfrac{\partial}{\partial\bar{\mu}}\left[\underset{(+)}{e\left(r_D,S\right)}\right]\bar{\mu} + \underset{(+)}{e(C,S)}\cdot\left[(\bar{\mu}/r_D) - \underset{(-)}{e(r_D,\bar{\mu})}\right] + \underset{(-)}{e\left(r_D,S\right)}\cdot \underset{(-)}{e\left(C,\bar{\mu}\right)} - \underset{(-)}{e\left(r_D,S\right)}\cdot \dfrac{\bar{\mu}}{\underset{(+)}{e\left(C,S\right)}}\cdot \dfrac{\partial}{\underset{(-)}{\partial\bar{\mu}}}\left[e\left(C,S\right)\right]}. \qquad (19)$$

Therefore, the bank’s perceived efficiency frontier has slope

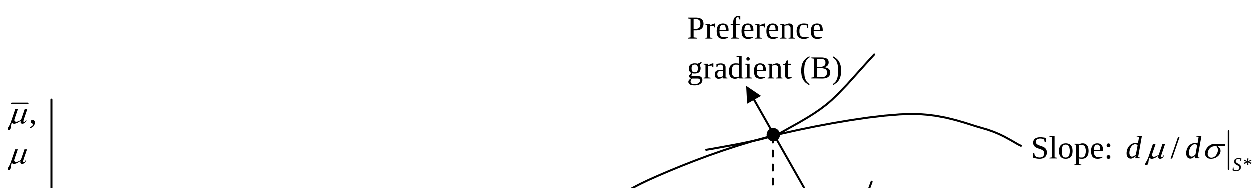
$$\left.\frac{d\mu}{d\sigma}\right|_{S^*} < 0 \text{ if } \begin{cases} \bar{\sigma} \text{ and } \bar{\mu} \to 0 & \text{(terms no. 1 and 4 of the numerator and denominator vanish)} \\ \bar{\sigma} \text{ and } \bar{\mu} \to \infty & \text{(term no. 4 of the numerator and denominator dominates since } \left|e\left(r_D S\right)\right| > \left|\partial/\partial\bar{\sigma}\left[e\left(r_D,S\right)\right]\right| \text{ and} \\ & \left|e\left(r_D,S\right)\right| > \left|\partial/\partial\bar{\mu}\left[e\left(r_D,S\right)\right]\right|. \end{cases} \qquad (20)$$

$$\left.\frac{d\mu}{d\sigma}\right|_{S^*} \lesseqgtr 0 \text{ otherwise.}$$

Figure 1 illustrates several efficiency frontiers to be discussed below. Note that μ and $\bar{\mu}$ as well as σ and $\bar{\sigma}$ are depicted on the same axis, respectively, reflecting the assumption that e.g. a low first-period value of $\bar{\sigma}$ tends to translate into a low third-period σ. While the slope $d\mu/d\sigma|_{S^*}$ is undetermined for intermediate values of $\bar{\mu}$ and $\bar{\sigma}$, it must be positive somewhere to avoid banks’ systematically attaining negative rates of return. Clearly, it is not constant but crucially depends on the relative size of the derivatives of $e\left(r_D,S\right)$ and $e\left(C,S\right)$ w.r.t. $\bar{\sigma}$ and $\bar{\mu}$ [see equation (19)].

Conclusion 3: Due to its lagged responses to shocks in expected rate of return and volatility, the bank induces an endogenous efficient frontier whose slope importantly depends e.g. on how the elasticity of the cost of refinancing w.r.t. the level of solvency reacts to an increase in volatility.

Figure 1: Endogenous tradeoffs in (μ,σ)-space



σ_I^*

4.2 Effects of solvency regulation on the efficiency frontier

The objectives of solvency regulation differ from those of the bank, who by assumption seeks to be on the efficient (μ,σ)-frontier as given in (20) and depicted as $d\mu / d\sigma|_{S*}$ in Figure 1. Solvency regulation is designed to avoid the external costs caused by insolvencies described in Section 2. Its main instrument is capital requirements, based on the norms of the Basel Committee on Banking Supervision, an agency of the Bank for International Settlements. Two types of regulations are analyzed below. The first requires a certain equity-to-assets ratio, independent of the portfolio risk, *Basel I*, linking capital requirements only to asset classes and not to asset risk per se. The second type corresponds to *Basel II*, which fixes capital requirements for different asset classes as a function of the risk profile of individual assets.

4.2.1 Basel I

Basel I stipulates capital requirements as a function of total assets and separately for off-balance sheet positions. These requirements therefore are independent of risk. In terms of the model, *Basel I* amounts to the restrictions

$$e(C,\bar{\sigma})=0\,,\ e(C,\bar{\mu})=0\,,\ e(C,S)=0\,,\ \frac{\partial e(C,S)}{\partial\bar{\sigma}}=0\,,\ \frac{\partial e(C,S)}{\partial\bar{\mu}}=0\,. \tag{21}$$

The two last restrictions are implied by the fact that $e(C,S)$ is a constant. Inserting them in (19), one now obtains for the slope of the trade-off (subscript *I* denoting *Basel I*),

$$\left.\frac{d\mu}{d\sigma}\right|_I = \frac{\dfrac{\partial}{\partial\bar{\sigma}}\left[e(\overset{(+)}{r_D},S)\right]\bar{\sigma}}{\dfrac{\partial}{\partial\bar{\mu}}\left[e(\underset{(+)}{r_D},S)\right]\bar{\mu}} \quad \begin{cases} 0 \; if \;\; \bar{\mu}\to\infty \\ > otherwise \end{cases}$$

(22)

The second statement holds also for $\bar{\mu}\to 0$ because in that case, replacing numerator and denominator by Taylor approximations and evaluing them at $\bar{\mu}=0$ yields the ratio of the two elasticities in (22), which is positive.

The slope defined by eq. (22) differs from expression (19). In particular, when $\bar{\sigma}$ and $\bar{\mu}\to 0$, $\left. d\mu/d\sigma\right|_{S*}<0$ in (19), while $\left. d\mu/\partial\sigma\right|_I$ in (22) remains positive. Moreover, the restrictions (21) cause the numerator of eq. (19) to increase and its denominator to decrease [with the sole exception of the term containing $\partial e\left(C,S\right)/\partial\bar{\mu}$, which however is of second-order magnitude compared to $e\left(C,S\right)$ itself]. Thus, the slope of the *Basel I* frontier is positive greater than that of the unregulated frontier for intermediate values of $\bar{\mu}$, approaching (but never crossing) the latter (after all, regulation cannot increase the bank's feasible set).

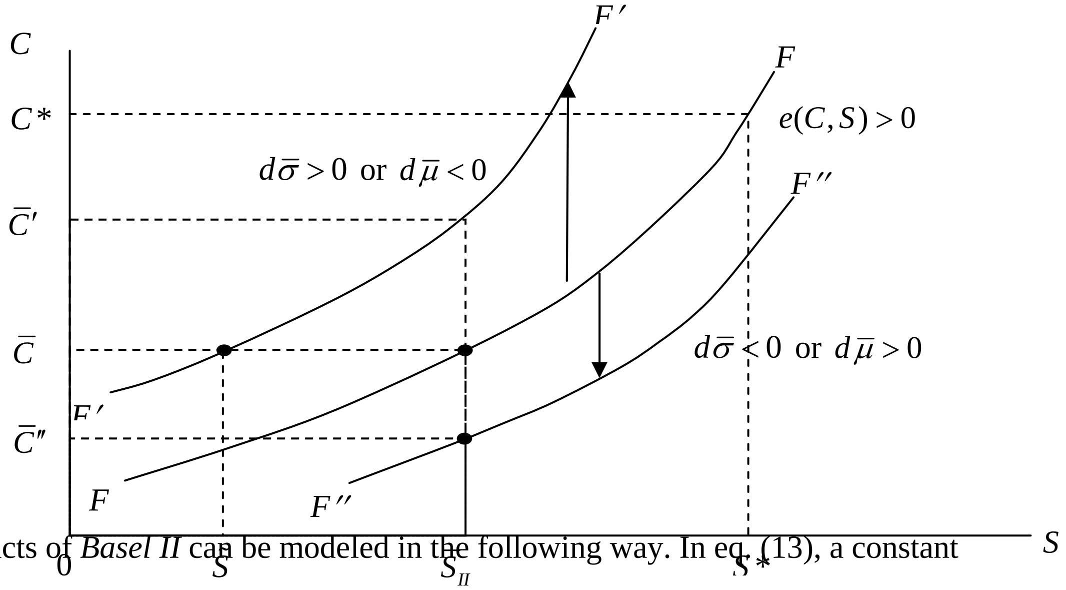
Figure 1 displays the pertinent efficiency frontier (marked with subscript *I*). For predicting optimal solutions, one needs two assumptions regarding the preferences of bank's management. First, while the RORAC objective of eq. (5) is defined in deterministic terms for simplicity, reflecting the interest of well-diversified owners of the bank, management typically exhibits risk aversion since it is much less diversified. Assuming imperfect governance, one is led to project their risk aversion into (μ,σ)-space, resulting in convex indifference curves. Second, homothecity is imposed in order to obtain sharper predictions. Under these assumptions, *Basel I* regulation induces the bank to be less conservative regardless of whether management is strongly risk-averse (type A) or weakly risk-averse (type B), i.e. ($\sigma_I^* > \sigma_{S*}^*$, $\sigma_I^{**} > \sigma_{S*}^{**}$).

Conclusion 4: Regulation of the *Basel I* type may induce banks to take a more risky position than they would on their own, thus having a counter-productive effect.

4.2.2 Basel II

Basel II allows a choice of approach for the calculation of capital requirements, viz. the Standardized Approach and the Internal Ratings-Based Approach. Whilst the first is based on *Basel I*, the second lets banks choose their probability of default, their percentage loss at default, and the maturity of their credits. Large institutions with average and below-average credit risks mostly choose the Internal Ratings-Based approach to save on capital despite its higher cost of implementation.

Figure 2: Implications of *Basel II* regulation



The impacts of *Basel II* can be modeled in the following way. In eq. (13), a constant $e(C,S)>0$ was assumed. This is represented by the progressively increasing slope of the curve labeled *F* of Figure 2. *Basel II* imposes a minimum degree of solvency, denoted by $\bar{S}_{II}$. Now let a shock $d\bar{\sigma}>0$ occur (volatility of returns has increased). Clearly, the locus labeled *F* shifts upward to F', indicating that a given capital $\bar{C}$ would now only suffice to guarantee a solvency level $\tilde{S}<\bar{S}_{II}$. Therefore, in order to satisfy the *Basel II* norm, a bank that just satisfied it initially would have to come up with the full additional amount of capital $\left(\bar{C}'-\bar{C}\right)$ to satisfy the solvency norm. A bank with excess solvency, symbolized by the combination $\left(S^*,C^*\right)$, would not have to react to the shock $d\bar{\sigma}>0$, however. The same conditional responses are predicted for a shock $d\bar{\mu}<0$, i.e. a drop in the mean return on investments.

Conversely, consider a shock $d\bar{\sigma}<0$, i.e. capital markets have become less volatile. This causes the locus *F* of Figure 2 to shift down to *F''*. Now $\bar{C}''<\bar{C}$ suffices to reach the prescribed solvency level, and the "marginal" bank that was at $\bar{S}_{II}$ initially can reduce capital by as much as $\left(\bar{C}-\bar{C}''\right)$. However, note that due to the convexity of the locus, the relaxation effect $\left(\bar{C}-\bar{C}''\right)$ is smaller than the tightening effect $\left(\bar{C}'-\bar{C}\right)$ for a shock of same absolute value. This of course holds true of $d\bar{\mu}>0$ as well.

In sum, one has the following set of conditional predictions for *Basel II*,

$$\frac{\partial}{\partial\bar{\sigma}}\left[e(C,S)\right]>0 \ \text{ if } \ d\bar{\sigma}>0 \ \text{ and } \ S=\bar{S}_{II}\,; \gtrless 0 \text{ if } d\bar{\sigma}>0 \ \text{ and } \ S>\bar{S}_{II}\,;$$

$$
\begin{aligned}
&\frac{\partial}{\partial \bar{\mu}}\left[e(C,S)\right] > 0 \quad \text{if} \quad d\bar{\mu} < 0 \text{ and } S = \bar{S}_{II} \\
&\qquad\qquad \gtrless 0 \quad \text{if} \quad d\bar{\mu} < 0 \quad \text{and} \quad S > \bar{S}_{II}; \\
&\frac{\partial}{\partial \bar{\sigma}}\left[e(C,S)\right] \lessgtr 0 \quad \text{if} \quad d\bar{\sigma} < 0; \\
&\frac{\partial}{\partial \bar{\mu}}\left[e(S,C)\right] \lessgtr 0 \quad \text{if} \quad d\bar{\mu} > 0.
\end{aligned}
\tag{23}
$$

In view of (19), one obtains

$$
\left.\frac{d\mu}{d\sigma}\right|_{II} > \left.\frac{d\mu}{d\sigma}\right|_{S*} \quad \text{if} \quad d\bar{\sigma} > 0,\ S = \bar{S}_{II}, \text{ and } \frac{\bar{\mu}}{r_D} \text{ large.} \tag{24}
$$

Figure 1 illustrates once more. At low values $\left(\bar{\mu}, \bar{\sigma}\right)$, the slope of the efficiency frontier induced by *Basel II* need not systematically differ from that induced by *Basel I*. The regulation being less stringent (at least by intent), the frontier runs higher than that of *Basel I*. However, even a strongly risk-averse management (preferences of type *A*) may still be induced to opt for a more risky allocation $\left(\sigma_{II}^{*} > \sigma_{S*}^{*}\right)$. According to (24), the frontier does run steeper than the one absent regulation at least for some banks when volatility of returns exogenously increases. Therefore, a less risk-averse management (preferences of type *B*) is predicted to pursue an investment policy that entails still higher volatility of returns than without regulation. The comparison with *Basel I* is ambiguous; however, the case $\sigma_{II}^{**} > \sigma_{I}^{**} > \sigma_{S*}^{**}$ (represented in Figure 1) cannot be excluded for some banks.

These banks just had satisfied the solvency norm initially, with expected returns exceeding refinancing cost r_D by a large amount. This combination of conditions is typical of an economic downturn – and it is precisely in this event when *Basel II* induces them to opt for a riskier position than they would on their own.

In sum, *Basel I* and *Basel II* are predicted to have similar effects in one respect. Both may induce at least some banks to opt for a more rather than less risky exposure than if they were optimizing free of the respective restraints. However, the two regulations differ in another respect. *Basel I* causes a "deformation" of the $\left(\mu, \sigma\right)$-frontier that depends on two things only, viz. (1) how strongly $e(r_D, S)$ reacts to change in $\bar{\mu}$ and $\bar{\sigma}$, respectively, and (2) the $\left(\mu, \sigma\right)$ position the bank is at initially. By way of contrast, *Basel II* leaves the basic trade-off [as given by eq. (19)] intact. Its “deformation effect” is limited to high values of $\bar{\mu}$ and a subset of banks.

Conclusion 5: At least for banks just compliant initially with the solvency norm and in some situations, *Basel II* may still cause banks to pursue a riskier investment policy than absent regulation, and possibly even riskier than under *Basel I*.

5 Summary and conclusion

The basic hypothesis of this paper states that banks seek to attain a certain solvency level that balances the advantage of lower refinancing cost against the disadvantage of tying capital that would yield higher returns in other uses. However, this solvency level is too low from a societal point of view because it neglects the fact that insolvency causes substantial external costs (Conclusion 1). The analysis proceeds to assume that banks maximize their rate of return of risk-adjusted capital (RORAC), which implies that the marginal benefit of a higher level of solvency is the lower cost of refinancing while its marginal cost consists of the extra capital to be allocated and return forgone (Conclusion 2). In the first period, two shocks occur, viz. an exogenous change in expected returns $\left(d\bar{\mu}\right)$ and in their volatility $\left(d\bar{\sigma}\right)$. These shocks induce lagged adjustments $\left(dS^*/d\bar{\mu}, dS^*/d\bar{\sigma}\right)$ during the second period. Net adjustment dS^* then triggers a reallocation of assets and liabilities and hence endogenous changes $d\mu$ and $d\sigma$ during the third period. This implies a perceived endogenous frontier in $\left(\mu,\sigma\right)$-space with slope $d\mu/d\sigma$. This slope is not a constant but depends importantly on e.g. how the elasticity of the cost of refinancing w.r.t. the level of solvency react to a change in volatility (Conclusion 3). The regulations imposed by *Basel I* are now shown to neglect several elasticities and their dependence on $\bar{\mu}$ and $\bar{\sigma}$, causing a modification of the risk-return frontier as perceived by regulated banks. This modification may induce them to take a more risky position than they would on their own (Conclusion 4). The implications of *Basel II* are more complex. Still, banks at the initially prescribed solvency level may react to an increase in volatility by again taking a more risky position than they would have otherwise, even more markedly so than under *Basel I* (Conclusion 5).

Both of these predicted adjustments may be considered counter-productive. However, it would be inappropriate to conclude that *Basel I* and *II* or even solvency regulation in general should be revoked. First, the model analyzed in this paper might be too simplistic; banks possibly pursue other objectives than just maximizing RORAC. Second, *Basel II* already constitutes an improvement over *Basel I* in that its counter-productive effect is limited to the (usually small) subset of banks that initially had just been compliant with the prescribed solvency level. And finally, assuming that solvency regulation does entail more benefit (in terms of external cost avoided) than cost (in terms of biasing banks' tradeoffs between μ and σ), one would have to find an alternative whose benefit-cost ratio beats that of *Basel I* and *II*. While this task is left for future analysis, the present work does call attention to likely shortcomings of current solvency regulation.

References

Basel Committee on Banking Supervision (1988): „International convergence of capital measurement and capital standards," Basel Capital Accord I, Bank for International Settlements.

Basel Committee on Banking Supervision (2005): „International convergence of capital measurement and capital standards: A revised framework," Basel Capital Accord II, Bank for International Settlements, Basel.

Bauer, W., and M. Ryser (2004): „Risk management strategies for banks," Journal of Banking & Finance, 28, 331-352.

Covitz, D.M., D. Hancock and M.L. Kwast (2004): "A reconsideration of the risk sensitivity of U.S. banking organization subordinated debt spreads: A sample selection approach", Economic Policy Review (Federal Reserve Bank of New York), 10, 73-92.

Diamond, D. W., and P. P. Dybvig (1983): „Bank runs, deposit insurance, and liquidity," Journal of Political Economy, 91, 401-419.

Flannery, M.J. and S.M. Sorescu (1996): "Evidence of bank market discipline in subordinated debenture yields: 1983-1991", Journal of Finance, 51, 1347-1377.

Furfine, C. H. (2003): „Interbank exposures: Quantifying the risk of contagion," Journal of Money, Credit, and Banking, 35, 111-128.

Goldberg, L. G., and S. C. Hudgins (1996): „Response of uninsured depositors to impending S&L failures: evidence of depositor discipline," The Quarterly Review of Economics and Finance, 36, 311-325.

Goldberg, L. G., and S. C. Hudgins (2002): „Depositor discipline and changing strategies for regulating thrift institutions," Journal of Financial Economics, 63, 263-274.

Jacklin, C. J., and S. Bhattacharya (1988): „Distinguishing panics and information-based bank runs: Welfare and policy implications," Journal of Political Economy, 96, 568-592.

Jensen, M. C., and W. H. Meckling (1976): „Theory of the Firm: managerial behavior, agency costs and ownership structure," Journal of Financial Economics, 3, 305-360.

John, K., A. Saunders, and L. W. Senbet (2000): „A theory of bank regulation and management compensation,” Review of Financial Studies, 13(1), Spring, 95-125.

Jordan, J. S. (2000): „Depositor discipline at failing banks,“ New England Economic Review, March/April, 15-28.

Kim, D., and A. Santomero (1988): „Risk in banking and capital regulation,” Journal of Finance, 43(5), Dec., 1219-1233

Koehn, M., and A. Santomero (1980): „Regulation of bank capital and portfolio risk,” Journal of Finance, 35(5), Dec., 1235-1244.

Lang, L. H. P., and R. M. Stulz (1992): „Contagion and competitive intra-industry effects of bankruptcy announcements,” Journal of Financial Economics, 32, 45-60.

Merton, R. C. (1974): „On the pricing of corporate debt: the risk structure of interest rates,” Journal of Finance, 29, 449-470.

Merton, R. C. (1977): „An analytical derivation of the cost of deposit insurance and loan guarantees,” Journal of Banking & Finance, 1, 3-11.

Park, S., and S. Peristiani (1998): „Market discipline by thrift depositors“, Journal of Money, Credit, and Banking, 30, 347-364.

Repullo, R. (2004): „Capital requirements, market power, and risk-taking banking,” Journal of financial Intermediation, 13, 156-182.

Rochet, J.-C. (1992): „Capital requirements and the behaviour of commercial banks,” European Economic Review, 36, 1137-1178.

Smith, C. W., and R. M. Stulz (1985): „The determinants of firms' hedging policies,“Journal of Financial and Quantitative Analysis, 20, 391-405.

Stulz, R. M. (1996): „Rethinking risk management,“ Journal of Applied Corporate Finance, 9, 8-24.