

Solvency Regulation of Banks and Insurers: A Two-Pronged Critique

Peter Zweifel¹

¹ University of Zurich, Switzerland

Correspondence: Peter Zweifel, Kreuth 371, A-9531 Bad Bleiberg, Austria. Tel: 43-4244-20319. E-mail: peter.zweifel@uzh.ch

Received: May 14, 2015

Accepted: June 4, 2015

Online Published: July 9, 2015

doi:10.5430/ijfr.v6n3p86

URL: <http://dx.doi.org/10.5430/ijfr.v6n3p86>

Abstract

This paper puts forth two models, one of a bank and one of an insurer. Their divisions responsible for investment (and risk underwriting, respectively) seek to maximize the expected risk-adjusted rate of return on capital (*RAROC*). For the bank, higher solvency lowers the cost of refinancing; for the insurer, it attracts more premium income. In both cases, however, higher solvency ties costly capital. Sequential decision making is tracked over three periods. In period 1, exogenous changes in expected returns and in volatility occur, each suggesting optimal adjustments in solvency. However, there can be only one adjustment, which in period 3 creates an endogenous trade-off between expected returns and volatility. This tradeoff is shown to involve parameters that differ between banks and insurers, calling for bank-specific and insurance-specific solvency regulation. In addition, *Basel I* and *II* (*Solvency I* and *II*, respectively) are shown to modify this tradeoff, inducing senior management of both banks and insurers to opt for higher volatility and lower solvency than otherwise. Therefore, existing solvency regulation can run counter their stated objective in both cases, which quite likely is also true of *Basel III* and *Solvency III*.

Keywords: regulation, banks, insurers, solvency, Basel I, II, III, Solvency I, II, III

1. Introduction

Risk-adjusted return on risk-adjusted capital (*RAROC*) has increasingly become the benchmark for assessing the performance and governance of banks' investment as well as insurers' underwriting and investment divisions. For banks, a higher solvency level has the benefit of lowering the cost of refinancing; for insurers, of enhancing demand and hence bolstering premium income. For both however, solvency ties costly capital. At the same time, public regulators are concerned about solvency to ensure the continuity of banking and insurance operations. This paper deals with the conflict between the optimization of solvency by banks and insurers themselves and solvency levels imposed by *Basel I* and *II* (*Solvency I* and *II*, respectively). It tracks the sequential decision making by both types of financial firms in the following way. In a first period, exogenous shocks on expected returns and volatility impinge on their investment (and underwriting) divisions, respectively. A typical cause could be that investments made in the previous period turn out to have a lower rate of return or a higher volatility than expected. In the second period, the divisions adjust their solvency levels in response to these shocks in ways predicted by maximization of *RAROC* and comparative-static analysis. In the third period, the divisions propose to senior management a rebalancing of asset position and underwriting portfolios in response to changed solvency levels. These endogenous adjustments entail a restructuring of portfolios resulting in new values for expected returns and volatility which define the slopes of internal efficiency frontiers. Finally, senior management chooses its optimum, taking into account its degree of risk aversion.

These endogenous efficiency frontiers are modified by solvency regulation such as *Basel I* and *II* and *Solvency I* and *II*, respectively. It will be argued that *Basel I* and *Solvency I* neglect the fact that for both banks and insurers, the relationship between solvency and their cost of refinancing depends on changes in expected returns and volatility that are exogenous to them. As to *Basel II*, and *Solvency II*, they address solvency directly but still fail to take into account that for banks and insurers that initially just met the prescribed solvency level, the extra amount of risk capital needed to maintain it also depends on the market environment with its changes in expected returns and volatility. It will be shown that both *Basel I* and *II* (*Solvency I* and *II*, respectively) modify the slope of the efficiency frontiers as perceived by regulated banks and insurers. While one might expect that solvency regulation reduces the slope of these frontiers (thus inducing senior management to opt for lower expected returns and lower volatility), it turns out that the opposite can well be the case. Indeed, through their neglect of parameters of importance to banks

and insurers themselves, *Basel I* and *II* as well as *Solvency I* and *II* may have the unexpected consequence of causing at least some of them to opt for a higher volatility in the rate of return on their assets than without it, causing regulation to miss its target. Moreover, the strength of this effect is shown to depend on bank-specific and insurance-specific parameters. Generally, it is likely to increase in the case of impending *Basel/Solvency III* regulation, without however attaining the same importance as in the case of *Basel I* and *Solvency I*.

These findings are important because the *Basel*-type regulations are predicated on the (unproven) belief that they induce banks to take on less risky positions, causing them to increase their solvency level. Likewise, the *Solvency*-type regulations take it for granted that insurers will opt for less risk exposure, again serving to increase their solvency level. Moreover, current solvency regulation tends to treat banks and insurers in the same way, although their business models differ, with clear implications for solvency regulation. While the critique contained in this paper is derived from two theoretical models, it is supported by some historical evidence; it may therefore provide a starting point for future regulation that is more attuned to the behavior and business models of banks and insurers.

This paper is structured as follows. Section 2 contains a review of the pertinent literature to conclude that solvency regulation of banks and insurers may serve to avoid negative externalities. In Section 3, a higher level of solvency is found to have two effects on banks' and insurers' decisions aiming to maximize *RAROC*. On the one hand, it serves to lower their cost of refinancing by lowering the interest rate paid on deposits and increasing premium income, respectively; on the other hand, it ties capital that would have other, more productive uses. This optimum is disturbed by exogenous shocks in terms of expected returns and volatility in the capital market (see period 1 of Figure 1). In period 2, the investment (and underwriting) divisions adjust the bank's and insurer's solvency levels to these shocks. These adjustments are derived in Section 4 using comparative-static analysis. However, there can be only one adjustment of solvency, which in period 3 acts like an exogenous change causing both banks and insurers to move along their endogenous efficiency frontiers. By performing comparative statics 'in reverse', the slopes of these frontiers are derived in Section 5. Reflecting the different business models, they depend on bank- and insurance-specific parameters, respectively. Senior managements are presented with this trade-off and make their choices taking account of their degrees of risk aversion (which may differ between banks and insurance companies). Regulations imposed by *Basel/Solvency I, II, and III* are introduced as parameter restrictions in Section 6 to show how the slopes of these efficiency frontiers are modified, likely inducing senior management of regulated banks and insurers to opt for a higher volatility and lower solvency than absent this regulation. However, this effect is again conditioned on parameters that differ between the two types of financial firms. A summary and conclusions follow in Section 7.

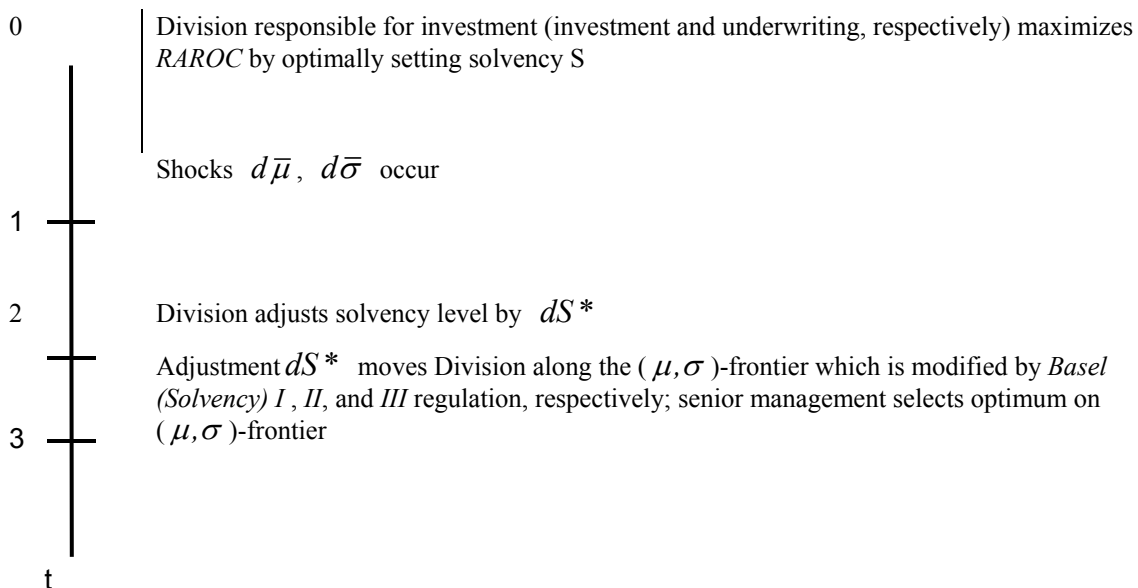


Figure 1. Timeline of the models

2. Literature Review

2.1 Solvency Regulation of Banks

Solvency regulation of banks has traditionally been justified by the view (sometimes held also by their shareholders) that their risk-taking is excessive. This view was challenged by the proponents of the Capital Asset Pricing Model (CAPM), who emphasized that for well-diversified investors, the solvency of a bank does not constitute a reasonable objective. Investors are concerned with expected profitability, adjusted for the degree to which the bank's profitability systematically varies with the capital market (the beta of the CAPM). By way of contrast, for little-diversified investors as well as ordinary consumers holding deposits, 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 (OPT) 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, among them depositors) of the bank [Merton (1974), Jensen and Meckling (1976), Merton (1977)]. Shareholders see the value of their stock 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]. Depositors in turn may have selected a low-solvency bank that pays a high rate of interest. Therefore, the cost of risk bearing can be said to be internal to both types of stakeholders.

From an efficiency point of view, the case for public intervention rests on external costs caused by insolvency. First, and most importantly, it 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, in addition to nonbank firms [Lang and Stulz (1992), Furfine (2003)]. Second, investors in the capital market at large may re-evaluate the estimated risk of insolvency. In response to revised estimates, they demand a higher rate of interest from solvent 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)].

2.2 Solvency Regulation of Insurers

Solvency regulation of insurers has been based on the claim of excessive risk-taking as in the case of banks (Cummins, 1988). Again however, internal and external costs of an insolvency need to be distinguished. There are well-diversified investors, for whom according to the CAPM the solvency of a particular insurer does not constitute a reasonable objective, while for little-diversified investors (among them, policyholders), the insurer's risk of insolvency is relevant. In terms of the OPT, the shareholders' put option is in the main financed by the insurer's policyholders [Cummins and Phillips (2001), Zweifel and Eisen (2011), ch. 6.3]. In the event of insolvency, internal costs fall on the shareholders of the insurance company. It deprives policyholders of their insurance coverage, causing them to lose at least part of their assets in the event of loss. Still, policyholders may have opted for a low-solvency, low-premium insurer.

Turning to external costs, some of the policyholders may be insurance companies involved in the coinsurance of large risks; therefore, the insolvent insurer may drive other financial institutions into insolvency (Furfine, 2003). Second, as in the case of banks, it may prompt investors in the capital market at large to revise their estimates of insurers' risk of insolvency, making them demand a higher rate of return. Policyholders may curtail their demand for coverage, resulting in a fall in premium income. Both reactions increase the cost of refinancing of solvent insurers. There is empirical research substantiating these claims [Cummins and Sommer (1996)].

It may be worthwhile to emphasize that externalities do not suffice to justify public regulation to ensure solvency. One would have to first examine whether the expected benefit of the intervention exceeds 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 since become '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. However, pointing to bank-specific problems of governance, Mülbert (2009) argues that prudential regulation of the *Basel I* and *II* type may even induce rather than prevent banking crises. This may be true of insurance as well since (European) insurers are subject to much the same solvency regulation as banks, *Solvency I* to *III* of the European Union being closely modeled after *Basel I* to *III*.

2.3 Contribution of the Present Paper

The research cited in the Sections 2.1 and 2.2 suggests that the solvency levels that are deemed optimal by individual banks and insurance companies are too low from a societal perspective. However, there is a major difference between the two financial firms. As mentioned in Section 2.1, the insolvency of a bank may trigger a bank run, with economy-wide and even global repercussions. A bank run is a possibility because (small) depositors can withdraw their funds at very low cost. In contrast, cancelling an insurance contract is fraught with considerable cost, making an insurance run a very unlikely event. For this reason alone, insurers constitute much less of a ‘systemic risk’ than do banks (Cummins and Weiss, 2011). Moreover, insurers provide protection against risk but do not provide liquidity as do banks. The insolvency of an insurer thus becomes relevant only in the event of loss (which usually is a rare event), whereas the insolvency of a bank blocks payments (which occur daily). Finally, costs of insolvency external to banks fall very quickly on central banks (and ultimately taxpayers) as lenders of last resort. In the case of an insolvent insurer, reinsurance provides at least partial protection to policyholders and other creditors (Davidson et al., 1992).

The present contribution differs from the existing literature in two ways. First, it clearly distinguishes between the earlier *Basel/Solvency I* and the more refined *Basel II/Solvency II* regulation, showing that the more recent variant may have unintended consequences only for a subset of banks and insurers rather than all of them. However, *Basel III/Solvency III* is found to likely increase this subset again. 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 [as in the case of Repullo (2004)] equilibria, here the bank’s and insurance company’s paths of adjustment from one optimum to the next are analyzed. Adjustment to exogenous shocks will be shown to be conditioned by solvency regulation of the *Basel/Solvency I* to *III* type. Third, its effects are shown to hinge on parameters that differ between banks and insurers, implying that solvency regulation needs to be tailored to the two types of financial firms rather than being uniform. In return, welfare implications will not be spelled out; rather, the fact that banks and insurers may be induced to act against the stated intentions of the regulator will be highlighted.

3. Optimal Solvency for Banks and Insurers

3.1 Optimal Solvency for a Bank’s Investment Division

Let a bank’s investment division (the Division henceforth) maximize the expected rate of risk-adjusted return on capital (*RAROC*) through its choice of solvency S . Solvency is defined as a value-at-risk (VaR) index, in keeping with *Basel II* [see Artzner et al. (1999) for a critique and Brandtner (2013) for a comparison between so-called coherent risk measures and VaR]. Whenever VaR increases, solvency can be said to decrease.

By assumption, senior management mandates the Division to act in a risk-neutral manner. This can be justified by noting that allowing risk aversion to affect decisions taken by employees of the Division would result in inconsistencies. Employee A (who is strongly risk averse) might turn down a client seeking to obtain funding for a project while employee B (who has more of a risk appetite) of the same bank would accept it. Senior management needs to avoid such inconsistencies. Therefore, the expectation operator is dropped as long as the Division is being analyzed; risk aversion will enter in period 3 when the bank’s senior management selects its preferred position on the (μ, σ) -efficiency frontier generated by the bank’s Division. A higher level of solvency S enables the bank (and hence the division) to obtain funds at a lower rate of interest paid on deposits r_D . Deposits are the only source of (re)financing considered in this paper; however, the global capital market as an alternative is characterized by investors who respond to a change in solvency in the same way. The bank is therefore confronted with refinancing cost according to

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

the arguments other than S are discussed in Section 4.1 below. The amount of risk-adjusted capital $C > 0$ increases with the solvency level S aimed at by the bank,

$$C = C(\cdot, S) \text{ with}$$

$$\frac{\partial}{\partial S} C(\cdot, S) > 0, \quad \frac{\partial^2 C}{\partial S^2} > 0. \quad (2)$$

RAROC is defined as the Division's net profit relative to risk-adjusted capital invested at the beginning of the period (for simplicity of notation, this difference in time is neglected) minus a hurdle rate r imposed by senior management. Net profit consists of net investment income $(\mu - r_D)D$, where μ denotes the rate of return on the capital market and D , deposits, and income $r_G C$ derived from investing solvency capital. The rate r_G reflects the fact that these funds must mainly be invested in gilt-edged securities, usually government bonds. Thus,

$$RAROC = \frac{(\mu - r_D(\cdot, S))D + r_G C(\cdot, S)}{C(\cdot, S)} - r.$$

Setting $r = r_G$ (although the hurdle rate usually exceeds the rate on government bonds), and neglecting operating expense and taxes for simplicity, *RAROC* can be expressed as a return on risk-adjusted capital,

$$RAROC = \frac{(\mu - r_D(\cdot, S))D + r_G C(\cdot, S) - r C(\cdot, S)}{C(\cdot, S)} = \frac{(\mu - r_D(\cdot, S))D}{C(\cdot, S)}. \quad (3)$$

Moreover, the volume of the business portfolio and hence D is kept constant during the three (rather short) periods. Then, maximization of *RAROC* in an initial period 0 leads to the following first-order condition for optimal solvency,

$$-\frac{\partial r_D[S^*]}{\partial S} - \frac{\mu - r_D[\cdot, S^*]}{C[\cdot, S^*]} \cdot \frac{\partial C[\cdot, S^*]}{\partial S} = 0. \quad (4)$$

The bracket notation points to the fact that the endogenous determinant S has to be evaluated at its optimal level. Equation (4) can be rewritten in terms of elasticities by multiplying through by S/r_D and defining $e(r_D, S) := (\partial r_D / \partial S) \cdot (S / r_D) < 0$ and $e(C, S) := (\partial C / \partial S) \cdot (S / C) > 0$ to become (for a different formulation, see Zweifel, 2013),

$$-e(r_D, S) - \frac{\mu - r_D}{r_D} \cdot e(C, S) = 0. \quad (5)$$

Equation (5) can be interpreted as follows. The Division needs to weigh the favorable marginal effect of increased solvency on the cost of refinancing (in percentage terms, first term of the equation) against its marginal downside effect (second term). 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 μ exceeds by far the bank's refinancing cost r_D . Note that because of $e(r_D, S) < 0$, it must be true that $\mu > r_D$ for an interior solution.

Interestingly, there is evidence suggesting that before the advent of solvency regulation, banks opted for higher solvency levels than those prescribed even by *Basel III*. Billings and Capie (2007), adjusting for hidden reserves, estimate true capital-asset ratios to have been as high as 8.12 percent during 1959-1967 among the five major UK banks, while *Basel III* mandates a (risk-weighted) common equity-asset ratio of 7 percent, which is equivalent to a capital-asset ratio of roughly 3.5 percent. Of course, changes in management compensation may also have contributed to this decline in capital-asset ratios (John et al., 2000). Still, solvency regulation of the *Basel* type may well have the scope of inducing banks to take on more rather than less risk, as will be found in Section 6.1.2 below.

3.2 Optimal Solvency for an Insurer's Division in Charge of Underwriting and Investment

The insurer's departments responsible for underwriting and investment are amalgamated into a single division, to avoid the complication of having two endogenous efficiency frontiers. As in the case of banks, the Division seeks to maximize the expected rate of return on risk-adjusted capital (*RAROC*) through its choice of solvency S , acting in risk-neutral manner. A higher level of solvency S enables the insurance company (and hence the Division) to obtain more funds through higher premium income [for empirical evidence, see Epermanis and Harrington (2006)]. Therefore, one has

$$r_G \text{ with } \frac{\partial}{\partial S} P(\cdot, S) > 0 \text{ and } \frac{\partial^2}{\partial S^2} P(\cdot, S) < 0; \quad (6)$$

the arguments other than S are discussed in Section 4.2 below. As in the case of a bank, the amount of risk-adjusted capital $C > 0$ increases progressively with the solvency level S aimed at by the insurer [see eq. (2)].

For an insurance company, *RAROC* reflects profits from two distinct activities, capital investment and risk underwriting. As to profits from investment activity, they derive first from solvency capital C (which is equated to risk-adjusted capital for simplicity) being invested at a rate of return r_G , reflecting the fact that insurers almost everywhere are legally obliged to hold government bonds. Second, current premium income makes funds available for investment according to the so-called funds-generating factor k (Cummins and Phillips, 2001). The higher k , the longer the lag between premiums received and claims paid. These funds can be invested at the rate of return μ prevailing on the capital market.

The insurer also derives profit from risk underwriting. This component is simply given by the difference between premium income P and losses paid L and abstracting from operating costs and taxes again, *RAROC* can be expressed as follows,

$$\begin{aligned} RAROC &= \frac{r_G \cdot C(\cdot, S) + k\mu \cdot P(\cdot, S) + P(\cdot, S) - L}{C(\cdot, S)} - r \\ &= \frac{(1 + k\mu)P(\cdot, S) - L}{C(\cdot, S)}. \end{aligned} \quad (7)$$

The second equality reflects the assumption $r_G = r$, as in Section 3.1.

Writing R for short again, one has the following first-order condition for optimal solvency, where $e(P, S) := (\partial P / \partial S)(S / P)$ denotes the elasticity of premium income and solvency capital w.r.t. the solvency level,

$$\begin{aligned} \frac{dR}{dS} &= \frac{(1 + k\mu)[\partial P / \partial S \cdot C - P \cdot \partial C / \partial S]}{C^2} \\ &= (1 + k\mu)[e(P, S) \frac{P}{C \cdot S} - e(C, S) \cdot \frac{P}{C \cdot S}] = 0 \end{aligned}$$

and hence

$$\frac{dR}{dS} = e(P, S) - e(C, S) = 0. \tag{8}$$

Note that losses L are exogenous to solvency, a reasonable assumption. By eqs. (6) and (2), both elasticities are positive, justifying neglect of boundary solutions ($S^* = 0$ in particular).

Equation (8) can be interpreted as follows. Since $e(P, S) > 0$, the first term represents the marginal benefit of increased solvency in percentage terms. The Division needs to weigh this marginal return against the marginal cost of solvency, which is given by $e(C, S) > 0$, reflecting the capital needed for a higher solvency level.

4. Adjustment of Solvency to Exogenous Shocks

4.1 Adjustment by Banks

During the first period, exogenous shocks impinging on rates of return ($d\bar{\mu}$) and volatility of returns ($d\bar{\sigma}$) occur (see Figure 1 again). To derive the Division’s optimal adjustments of the solvency level, the assumptions listed in Table 1 are introduced. The signs of the second-order derivatives (B6 to B9) are of particular importance. While necessary for the comparative-static analysis performed in Appendix A, they show that the relationship between solvency and the cost of refinancing as well as the amount of risk capital needed depends on conditions prevailing in the capital market.

As shown in Appendix A, optimal adjustment of the solvency level S^* to a shock $d\bar{\mu} > 0$ in expected returns is qualitatively given by

$$\text{sgn} \left[\frac{dS^*}{d\bar{\mu}} \right] = \text{sgn} \left[\frac{\partial^2 R}{\partial S \partial \bar{\mu}} \right] = \text{sgn} \left[-\frac{\partial^2 r_D}{\partial S \partial \bar{\mu}} + \frac{1}{\mu - r_D} \left(1 - \frac{\partial r_D}{\partial \bar{\mu}} \right) \frac{\partial r_D}{\partial S} - \frac{1}{C} \cdot \frac{\partial C}{\partial \bar{\mu}} \cdot \frac{\partial r_D}{\partial S} - \frac{\mu - r_D}{C} \cdot \frac{\partial^2 C}{\partial S \partial \bar{\mu}} \right]; \tag{9}$$

the terms are signed using assumptions B2 to B9. Therefore, one obtains

$$\frac{dS^*}{d\bar{\sigma}} \begin{cases} < 0 & \text{if } \mu - r_D \rightarrow 0; \\ > 0 & \text{if } \mu - r_D \gg 0. \end{cases} \tag{10}$$

These results are intuitive. If the bank’s margin $\mu - r_D$ is extremely small [note that the pertinent multiplier $(1 - \partial r_D / \partial \bar{\mu})$ is bounded by (0,1)], the Division’s need to preserve costly capital becomes of overriding importance, causing it to reduce its solvency level in response to an exogenous increase in expected returns. However, when the margin becomes larger, less capital is needed to attain a given solvency level. This permits to actually increase the solvency level. Thus, $dS^* / d\bar{\mu} > 0$ is considered the normal response.

Now consider a shock $d\bar{\sigma} > 0$ (again, details are given in Appendix A),

$$\text{sgn} \left[\frac{dS^*}{d\bar{\sigma}} \right] = \text{sgn} \left[\frac{\partial^2 R}{\partial S \partial \bar{\sigma}} \right] = \text{sgn} \left[-\frac{\partial^2 r_D}{\partial S \partial \bar{\sigma}} - \frac{1}{\mu - r_D} \cdot \frac{\partial r_D}{\partial \bar{\sigma}} \cdot \frac{\partial r_D}{\partial S} - \frac{1}{C} \cdot \frac{dC}{d\bar{\sigma}} \cdot \frac{\partial r_D}{\partial S} - \frac{\mu - r_D}{C} \cdot \frac{\partial^2 C}{\partial S \partial \bar{\sigma}} \right]. \tag{11}$$

Using assumptions B2 to B9 once more, one obtains

$$\frac{dS^*}{d\bar{\sigma}} \begin{cases} > 0 & \text{if } \mu - r_D \text{ small}; \\ < 0 & \text{if } \mu - r_D \rightarrow \infty. \end{cases} \tag{12}$$

Again, the results are intuitive. An exogenous increase in volatility of returns makes refinancing more costly; to counteract this effect, it is appropriate to increase the solvency level if the margin is small. Note that ‘small’ does not imply ‘close to zero’ in this case because the relevant term contains $\partial r_D / \partial \bar{\sigma}$ and $\partial r_D / \partial S$, which are both first-order. Thus, $dS^* / d\bar{\sigma} > 0$ can be regarded the normal response. Yet, in the presence of very high margins, the opportunity cost of increased solvency becomes excessive, motivating a decrease in solvency.

Table 1. Assumptions of the bank model

B1:	$\mu = \bar{\mu} + \hat{\mu};$ $\sigma = \bar{\sigma} + \hat{\sigma}.$	Returns and volatility (μ, σ) are additive in an exogenous $(\bar{\mu}, \bar{\sigma})$ component determined on the capital market and an endogenous one
B2:	$\partial C / \partial \bar{\mu} < 0.$	The higher returns on the capital market, the less risk capital is needed to attain a given solvency level. A positive shock on returns makes positive net values of the bank more likely, therefore reducing the need for risk capital.
B3:	$\partial C / \partial \bar{\sigma} > 0.$	The higher volatility on the capital market, the more risk capital is needed to attain a given solvency level. Positive net values of the bank are less likely, and this must be counteracted by more risk capital.
B4:	$0 < \frac{\partial r_D}{\partial \bar{\mu}} < 1.$	The rate of interest paid on deposits reacts to an exogenous increase of returns less than proportionally. Otherwise, the condition $\hat{\mu} > r_D$ for an interior optimum [see eq. (4) again] would sooner or later be violated.
B5:	$\frac{\partial r_D}{\partial \bar{\sigma}} > 0.$	With increased volatility in the market, the bank must offer better conditions to depositors as well.
B6:	$\frac{\partial^2 r_D}{\partial S \partial \bar{\mu}} < 0.$	According to B4, the bank must increase its interest rate on deposits when market conditions become more favorable. However, it can afford to adjust to a lesser degree if its solvency level is high.
B7:	$\frac{\partial^2 r_D}{\partial S \partial \bar{\sigma}} < 0.$	According to B5, the bank must follow the market with its interest paid on deposits. However, it can again afford to adjust to a lesser degree if its solvency level is high.
	$\frac{\partial^2 r_D}{\partial S \partial \bar{\mu}} < 0.$	The inequality derives from the fact that by B4, $\partial r_D / \partial \bar{\mu}$ is bounded, while $\partial r_D / \partial \bar{\sigma}$ is not.
B8:	$\frac{\partial^2 C}{\partial S \partial \bar{\mu}} < 0.$	A higher solvency level calls for more risk capital but to a lesser degree if higher market returns prevail, making positive net values of the bank more likely.
B9:	$\frac{\partial^2 C}{\partial S \partial \bar{\sigma}} > 0.$	A higher solvency level calls for more risk capital, especially when market volatility is high, making positive net values of the bank less likely.

4.2 Adjustment by Insurers

In the first period, exogenous shocks impinging on rates of return ($d\bar{\mu}$) and volatility of returns ($d\bar{\sigma}$) occur (see Figure 1 again). As shown in Appendix B, the optimal adjustment of the solvency level S^* to a shock in expected returns $d\bar{\mu} > 0$ is qualitatively given by

$$\text{sgn} \left[\frac{\partial^2 R}{\partial S \partial \bar{\mu}} \right] = \text{sgn} \left[\frac{\partial^2 P}{\partial S \partial \bar{\mu}} \cdot \frac{S}{P} + \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial P / \partial \bar{\mu}}{P^2} \right) - \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial C / \partial \bar{\mu}}{C^2} \right) - \frac{\partial^2 C}{\partial S \partial \bar{\mu}} \cdot \frac{S}{C} \right]$$

and hence (with signs according to assumptions I1, I3, and B8),

$$\begin{aligned} \frac{dS^*}{d\bar{\mu}} &\rightarrow 0 \text{ if } S \rightarrow 0; \\ \frac{dS^*}{d\bar{\mu}} &> 0 \text{ if } C \text{ large, since } C^2 \rightarrow 0 \text{ faster than } C \rightarrow 0; \\ \frac{dS^*}{d\bar{\mu}} &< 0 \text{ otherwise.} \end{aligned} \tag{13}$$

These results are intuitive. In a situation where the solvency level is very low to begin with, its adjustment in response to increased returns in the capital market does not matter. However, with solvency capital C large, the opportunity cost of an increased solvency level is small, leading the Division to propose such an increase to senior management with the aim of boosting premium income. Since most insurers have excessive solvency capital [Nakada et al. (1999)], $dS^* / d\bar{\mu} > 0$ is considered the normal response.

Now consider a shock $d\bar{\sigma} > 0$ (again, details are given in Appendix B)

$$\text{sgn} \left[\frac{\partial^2 R}{\partial S \partial \bar{\sigma}} \right] = \text{sgn} \left[\frac{\partial^2 P}{\partial S \partial \bar{\sigma}} \cdot \frac{S}{P} + \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial P / \partial \bar{\sigma}}{P^2} \right) - \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial C / \partial \bar{\sigma}}{C^2} \right) - \frac{\partial^2 C}{\partial S \partial \bar{\sigma}} \cdot \frac{S}{C} \right], \tag{14}$$

implying (with terms signed according to assumptions I2, I4, and B9),

$$\begin{aligned} \frac{dS^*}{d\bar{\sigma}} &\rightarrow 0 \text{ if } S \rightarrow 0; \\ \frac{dS^*}{d\bar{\sigma}} &> 0 \text{ if } C \text{ small since } C^2 \rightarrow 0 \text{ faster than } C \rightarrow 0; \\ \frac{dS^*}{d\bar{\sigma}} &< 0 \text{ if } P \text{ small since } P^2 \rightarrow 0 \text{ faster than } P \rightarrow 0; \\ \frac{dS^*}{d\bar{\sigma}} &< 0 \text{ otherwise.} \end{aligned} \tag{15}$$

Again, the predictions are intuitive. If solvency is very low initially, adjusting it to an exogenous increase in the volatility of returns makes no difference. If solvency capital is scarce, it makes sense to increase the solvency level in order to attract more capital through premium income. On the other hand, in a situation where premium income is quite low to begin with, increasing it through higher solvency has little leverage; therefore, preserving scarce capital takes precedence, suggesting a decrease in solvency.

Table 2. Assumptions of the insurer model (for derivatives of capital C , see Table 1)

I1:	$\frac{\partial P}{\partial \bar{\mu}} < 0.$	The (present value of) premium income depends negatively on the rate of return attainable on the capital market because policyholders now have more favorable investment alternatives.
I2:	$\frac{\partial P}{\partial \bar{\sigma}} > 0.$	The (present value of) premium income depends positively on the volatility of returns on the capital market because the insurer now offers a comparatively safe investment alternative to risk-averse policyholders.

- I3: $\frac{\partial^2 P}{\partial S \partial \bar{\mu}} > 0$. While a higher rate of return on the capital market depresses premium income (see I1), this effect is weakened if the solvency level is high.
- I4: $\frac{\partial^2 P}{\partial S \partial \bar{\sigma}} > 0$. Higher volatility on the capital market serves to increase premium income (see I2); this effect is reinforced if the solvency level is high.

5. Determination of Endogenous Efficiency Frontiers

5.1 The Endogenous Efficiency Frontier of the Bank

In the third period, the bank inherits a net adjustment of solvency dS^* from the second period that is the result of responses to the shocks $(d\bar{\mu}, d\bar{\sigma})$ that occurred in the first period. The Division now proceeds to adjust the endogenous components $\hat{\mu}$ and $\hat{\sigma}$. Performing comparative statics ‘in reverse’, optimal adjustments can be predicted using eqs. (9) and (11), respectively, with dS^* now assuming the role of an exogenous shock. By reshuffling the bank’s assets, the Division therefore effects changes $d\hat{\mu}$ and $d\hat{\sigma}$, creating an endogenous efficiency frontier with slope $d\hat{\mu}/d\hat{\sigma}$, on which senior management will proceed to choose the optimum (see Section 6.1). This slope can be obtained by dividing (11) by (9), yielding

$$\left. \frac{d\hat{\mu}}{d\hat{\sigma}} \right|_{S^*} = \frac{-\frac{\partial^2 r_D}{\partial S \partial \bar{\sigma}} \frac{(-)}{\mu - r_D} \frac{(+)}{\partial \bar{\sigma}} \frac{(+)}{\partial S} \frac{(-)}{\partial r_D} \frac{(-)}{C} \frac{(+)}{\partial \bar{\sigma}} \frac{(-)}{\partial S} \frac{(+)}{\mu - r_D} \frac{(+)}{\partial S \partial \bar{\sigma}} \frac{\partial^2 C}{\partial S \partial \bar{\sigma}}}{-\frac{\partial^2 r_D}{\partial S \partial \bar{\mu}} \frac{(-)}{\mu - r_D} \frac{(+)}{\partial \bar{\mu}} \left(1 - \frac{\partial r_D}{\partial \bar{\mu}} \right) \frac{(+)}{\partial S} \frac{(-)}{\partial r_D} \frac{(-)}{C} \frac{(+)}{\partial \bar{\mu}} \frac{(-)}{\partial S} \frac{(-)}{\partial r_D} \frac{(+)}{\mu - r_D} \frac{(+)}{\partial S \partial \bar{\mu}} \frac{\partial^2 C}{\partial S \partial \bar{\mu}}}. \tag{16}$$

The sign of eq. (16) is negative both if $(\mu - r_D) \rightarrow 0$ and if $(\mu - r_D) \rightarrow \infty$ in view of eqs. (10) and (12). In Figure 2, these extreme cases are shown as dashed segments for completeness. However, with $dS^*/d\bar{\mu} > 0$ and $dS^*/d\bar{\sigma}$ constituting the normal responses [see the discussion below eqs. (10) and (12)], the slope of the endogenous efficiency frontier is positive for intermediate values of $(\mu - r_D)$ and hence μ by assumptions B1 and B4. Moreover, a negatively sloped internally perceived efficiency frontier in (μ, σ) -space would contradict daily experience on capital markets $(d\mu/d\sigma > 0)$. A crucial result is that the slope defined in eq. (16) depends not only on observable parameters such as (μ, r_D) and first-order effects the regulator likely is aware of such as $(\partial C/\partial \bar{\mu}, \partial C/\partial \bar{\sigma}, \partial r_D/\partial S)$ but also terms such as $\partial^2 C/\partial S \partial \bar{\mu}$ and $\partial^2 C/\partial S \partial \bar{\sigma}$ which indicate that the relationship between required risk capital and solvency depends on conditions on the capital market (see assumptions B8 and B9 again).

Figure 2 shows three endogenous efficiency frontiers. Note that μ and $\bar{\mu}$ as well as σ and $\bar{\sigma}$ are depicted on the same axis, reflecting the assumption that e.g. a low first-period value of $\bar{\sigma}$ tends to translate into a low third-period σ . The first frontier (labeled S^*) holds prior to the influence of regulation. The two other frontiers (labeled I and II , respectively) are modified by *Basel I* and *Basel II* regulation in ways to be discussed in Section 6.1 below.

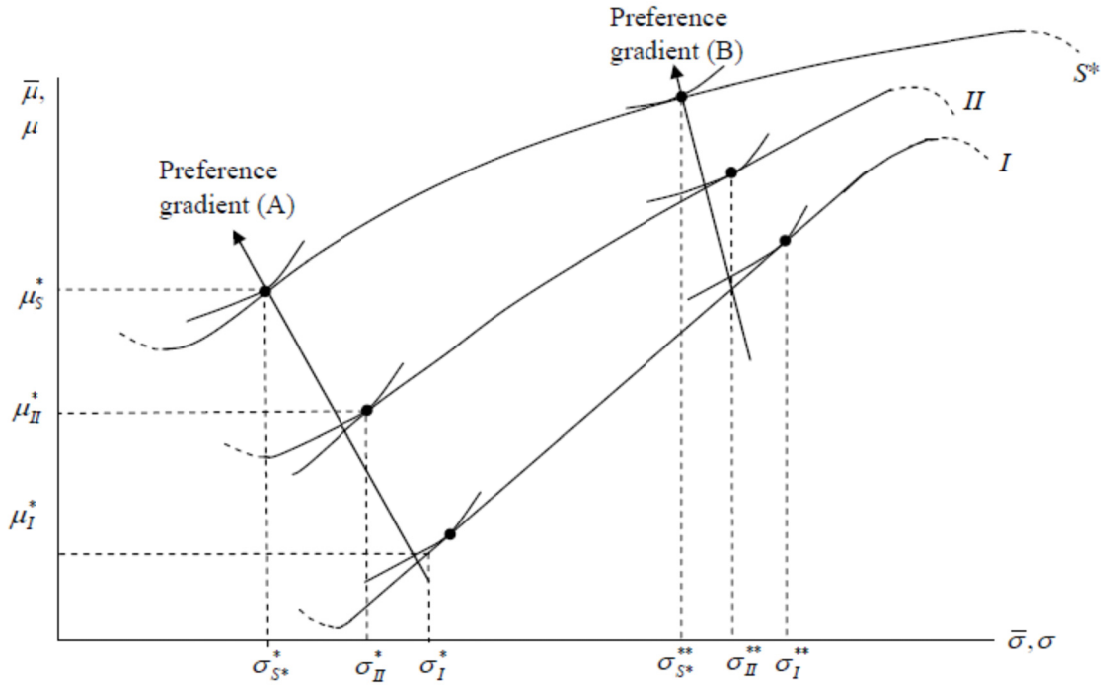


Figure 2. Endogenous efficiency frontiers in (μ, σ) -space

5.2 The Endogenous Efficiency Frontier of the Insurer

In the third period, the insurer inherits a net adjustment of solvency dS^* from the second period the same way as the bank. Again performing comparative statics in reverse, one can derive optimal endogenous adjustments $d\hat{\mu}/dS$ and $d\hat{\sigma}/dS$, respectively. The Division effects these changes by reshuffling the portfolios of assets and the risks, creating an endogenous efficiency frontier with slope $d\hat{\mu}/d\hat{\sigma}$. The slope of this frontier can be obtained by dividing (14) by (13), yielding

$$\left. \frac{d\hat{\mu}}{d\hat{\sigma}} \right|_{S^*} = \frac{\frac{\partial^2 P}{\partial S \partial \bar{\sigma}} \cdot \frac{S}{P} + \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial P / \partial \bar{\sigma}}{P^2} \right) - \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial C / \partial \bar{\sigma}}{C^2} \right) - \frac{\partial^2 C}{\partial S \partial \bar{\sigma}} \cdot \frac{S}{C}}{\frac{\partial^2 P}{\partial S \partial \bar{\mu}} \cdot \frac{S}{P} + \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial P / \partial \bar{\mu}}{P^2} \right) - \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial C / \partial \bar{\mu}}{C^2} \right) - \frac{\partial^2 C}{\partial S \partial \bar{\mu}} \cdot \frac{S}{C}} > 0. \quad (17)$$

In principle, the sign of eq. (17) is indeterminate; however, as in the case of the bank, a positive sign is assumed in the following, resulting in a very similar efficiency frontier (see Figure 2). Since the normal response makes the denominator of eq. (17) positive, its numerator must be positive as well, an implication that will be of relevance in Section 6.1.2 below. Again, the slope defined in eq. (17) depends not only on easily observable parameters such as C, P, S and first-order effects the regulator likely is aware of [such as $\partial C / \partial \bar{\mu}, \partial C / \partial \bar{\sigma}$], but also terms such as $\partial^2 C / \partial S \partial \bar{\mu}$ and $\partial^2 C / \partial S \partial \bar{\sigma}$ which indicate that the relationship between solvency and solvency capital depends on conditions on the capital market (see assumptions I3 and I4 again).

As in the bank model, the first frontier (labeled S^*) holds prior to the influence of regulation. The two other frontiers (labeled I and II , respectively) are modified by *Solvency I* regulation in ways to be discussed in Section 6.1.2 below.

Conclusion 1: Due to its responses to shocks in expected rate of return and volatility in the process of sequential adjustment, the investment divisions of banks (and underwriting and investment divisions of insurers, respectively) induce an endogenous efficient frontier, whose slope depends on the changing relationship between solvency and capital.

There is some historical evidence supporting this conclusion. After adjusting for hidden reserves, Billings and Capie (2007) find that capital-asset ratios of the five major UK banks were particularly high during 1942-1946, when they had to finance the war effort. Their explanation is "...that much of their lending to government was in the form of marketable securities, which have generated *exposure to fluctuations in market prices*" (p. 152, emphasis added). Banks apparently realized that their efficiency frontiers depended on changes in the relationship between solvency and risk capital due to changing market conditions during the second World War.

6. Effects of Solvency Regulation on the Efficiency Frontier

6.1 Basel/Solvency I Regulation

6.1.1 Effects on Banks

Solvency regulation is designed to avoid the external costs caused by insolvencies described in Section 2.1. Its objectives of differ from those of the bank, who by assumption seeks to be on the efficient (μ, σ) -frontier as given

in (9) and depicted as $d\mu/d\sigma|_{S^*}$ in Figure 2. *Basel I* stipulates capital requirements as a function of risk-weighted assets and separately for off-balance sheet positions (Basel Committee on Banking Supervision, 1988). Its focus is on the relationship between solvency and capital. By defining four asset classes with fixed weights, *Basel I* imposes a fixed relationship between solvency capital C and solvency S (see the locus B of Figure 3 below). It therefore does not allow banks to react to changes in market conditions affecting the risk characteristic of assets. In terms of the model, this neglect amounts to the restrictions

$$\frac{\partial^2 C}{\partial S \partial \bar{\mu}} = 0, \quad \frac{\partial^2 C}{\partial S \partial \bar{\sigma}} = 0. \quad (18)$$

Inserting this in eq. (16), one immediately sees that the numerator increases while the denominator increases. One therefore obtains for the slope of the efficiency frontier (subscript I denoting *Basel I*, subscript S^* , the bank's own solvency management),

$$\frac{d\hat{\mu}}{d\hat{\sigma}}|_I \geq \frac{d\bar{\mu}}{dS^*}|_{S^*}. \quad (19)$$

The *Basel I* frontier therefore runs steeper than the bank's own S^* frontier, approaching but never crossing it for high values of μ because regulation cannot increase the bank's feasible set (see Figure 2).

One might argue that the bank can choose to act in accordance with parameters it knows to be of importance, contrary to the regulator's decision rule. This amounts to neglecting the restrictions stated in eq. (18). However, as emphasized by Power (2004, ch. 7), managers are responsibility-averse, leading them to use regulatory decision rules as a convenient justification of their own actions. For example, let there be a second-period upward adjustment in solvency indicating that the bank should move away from the origin on the efficiency frontier. With the flat endogenous efficiency frontier S^* auf Figure 2 in view, the investment division would propose to accept a substantial increase in volatility whereas based on the steeper efficiency frontier induced by *Basel I*, the suggested increase is smaller. If the bank's senior management were to move along S^* , it could be criticized by the regulator for taking on an excessive amount of risk. This threat causes a responsibility-averse management to adopt the restrictions of eq. (18) and view the steeper *Basel I* efficiency frontier as the relevant one.

For predicting optimal solutions, one needs two assumptions regarding the preferences of bank's senior management. The first is the assumption of risk aversion on the part of senior management. In principle, fully diversified owners would like management to act in risk-neutral manner (see the discussion in Section 2.1). Accepting a great deal of risk allows them to benefit from an increase in the value of their shares [which have properties of a call option, see

e.g. Zweifel and Eisen (2012), ch. 4.3]. Managers however, being less than perfectly diversified with their human capital tied to the employing firm, have reason to be risk-averse. Given imperfect governance by (dispersed) ownership, management can follow its preferences at least to some extent (Shrieves and Dahl, 1992). Second, homotheticity of risk preferences is imposed in order to obtain sharper predictions. Under these assumptions, *Basel I* regulation induces the bank to be less conservative regardless of degree of risk aversion (type A or B in Figure 2; see the movements from $\sigma_{S^*}^*$ to σ_I^* and from $\sigma_{S^*}^{**}$ to σ_I^{**} , respectively).

6.1.2 Effects on Insurers

Solvency I, being modeled closely after *Basel I* by the European Union, subjects (European) insurance companies to the same type of regulation as banks (see Section 6.1.1). In terms of eq. (17), which defines the slope of the insurer's efficiency frontier, this amounts to the same restrictions as stated in eq. (18). Inserting these restrictions in eq. (17), one immediately sees that the numerator increases while the denominator decreases. One therefore obtains for the slope of the efficiency frontier (subscript *I* denoting *Solvency I*),

$$\left. \frac{d\hat{\mu}}{d\hat{\sigma}} \right|_I > \left. \frac{d\hat{\mu}}{d\hat{\sigma}} \right|_{S^*} \quad (20)$$

The insurer's *Solvency I* frontier therefore runs steeper than the original *S** frontier, as in the case of banks. Introducing the same homotheticity assumption concerning risk preferences of senior management, one has

Conclusion 2: Regulation of the *Basel/Solvency I* type is predicted to induce banks and insurers to take a more risky position than they would on their own, running counter stated objectives of solvency regulation.

One could argue that according to Figure 2, these increases in volatility go along with increases in expected returns; therefore, solvency as defined by *VaR* need not diminish. To assess this argument, one can use the lognormal distribution of returns as an approximation. Then, an increase of σ by x percent would have to be associated with an increase in μ by x percent for *VaR* to remain constant, implying unitary elasticity [in analogy to Cummins and Nye (1981)]. In Figure 2, the corresponding locus is a straight line through the origin with a 45 degree slope. However, much lower slopes (typically below 0.5) have been found in empirical research [see e.g. Woehrmann et al. (2004)]. Therefore, the predicted increases in σ indeed cause a reduction in solvency.

While Conclusion 2 is qualitatively the same for banks and insurance companies, the impacts of *Basel I* and *Solvency I* may well differ in magnitude, depending on parameters specific to the respective business models. For the bank, these parameters are [see eq. (16)],

- the margin between the rate of return μ achievable on the capital market and the rate r_D that must be paid on deposits. The higher this margin, the stronger the steepening effect of *Basel I* on the efficiency frontier;
- the amount of solvency capital C . The smaller C , the stronger the steepening effect of *Basel I* on the efficiency frontier.

By way of contrast, the parameters conditioning the effect of *Solvency I* on insurance companies are [see eq. (17)],

- the level of solvency S already achieved. The higher S , the stronger the steepening effect of *Basel I* on the efficiency frontier;
- the amount of solvency capital C . The smaller C , the stronger the steepening effect of *Basel I* on the efficiency frontier. While this influence is shared with the banks, its magnitude may still differ.

In sum, the 'deformation effect' of *Basel I* on banks is particularly marked when business conditions are favorable. In the case of insurers, this is true of *Solvency I* when their initial situation is one of high solvency. Both banks and insurance companies are affected strongly when their capital base is small. However, differentiation of eqs. (16) and (17) w.r.t. C reveals that the magnitude of the effect depends on a host of parameters that are specific to the respective business model [notably $(\mu - r_D)$ and $\partial r / \partial S$ in the case of banks, S and $\partial P / \partial S$ in the case of insurers) or that take on values that may well differ between the two types of financial firms ($\partial^2 C / \partial S \partial \bar{\mu}$ and $\partial^2 C / \partial S \partial \bar{\sigma}$). One is therefore led to conclude that solvency regulation needs to be tailored to the specific properties of banks and insurers.

6.2 Basel II/ Solvency II Regulation

Basel II/ Solvency II allows a choice of approach for the calculation of capital requirements, viz. the Standardized Approach and the Internal Ratings-Based Approach (Basel Committee on Banking Supervision, 2004). Whilst the first approach is based on Basel I, the second lets banks and insurance companies choose their probability of default, their percentage loss at default, and the maturity of their credits and liabilities. Large institutions with average and below-average risks mostly prefer the Internal Ratings-Based approach to save on capital despite its higher cost of implementation. In terms of the model, Basel II/SolvencyII permits banks and insurers to take all elements of eqs. (16) and (17), respectively into account, which amounts to a lifting of the restrictions stated in eq. (18) as long as the constraint regarding the solvency level is not binding.

To show the consequences of Basel II/Solvency II, assume that the bank (the insurance company, respectively) has opted for the more flexible Internal-Ratings-Based approach. Taken together, the rules promulgated by the Basel Committee on Banking Supervision (2004, especially para. 40 to 44, which serve as the model for Solvency II) establish a relationship between a targeted solvency level and required risk capital. This relationship (which cannot be modified once an internal rating model is selected) is depicted by the locus labeled *F* in Figure 3. It has increasing slope to reflect decreasing marginal returns to risk capital as stated in eq. (2). Locus *F* runs below locus *B* of Basel I/Solvency I, reflecting that Basel II/Solvency II allows banks and insurers to save on risk capital. The other difference is that Basel II/Solvency II imposes a minimum degree of solvency, which is denoted by \tilde{S}_{II} .

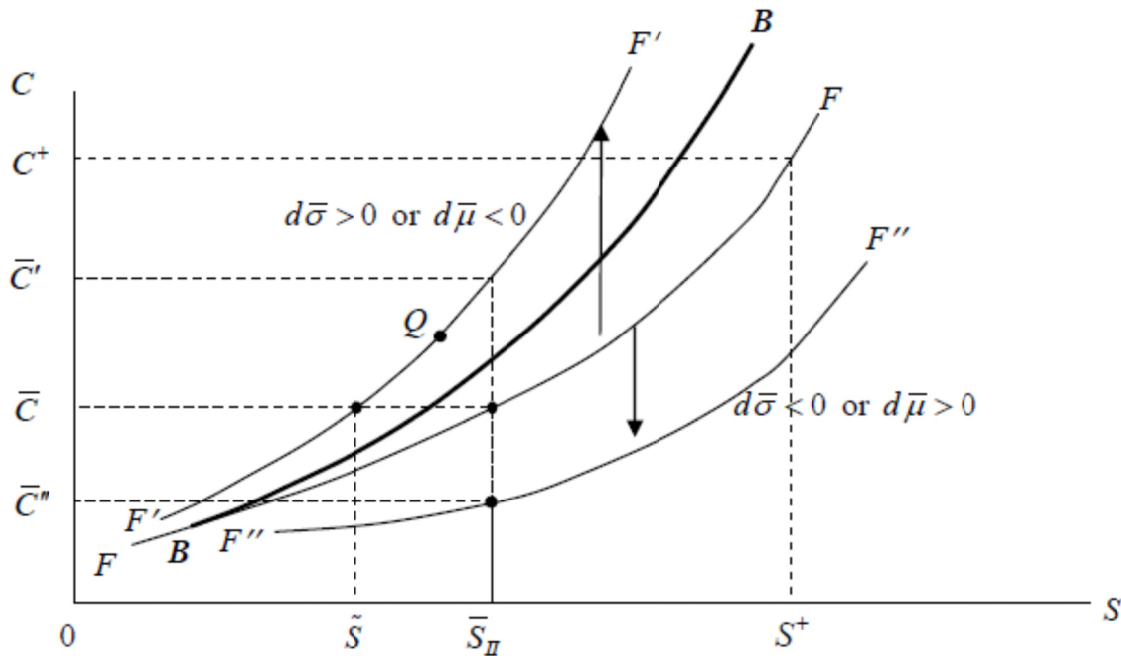


Figure 3. Consequences of Basel II/Solvency II regulation

Now let a shock $d\bar{\sigma} > 0$ occur (volatility of returns in the capital market has increased). In keeping with assumption B9, this corresponds to a steepening of the locus *F*, resulting in *F'* from the firm's point of view. It indicates that a given capital \bar{C} would now only suffice to guarantee a solvency level $\tilde{S} < \bar{S}_{II}$. Therefore, in order to keep to the Basel II/Solvency II norm, a bank or an insurance company that just satisfied it initially would have to come up with the full additional amount of capital $(\bar{C}' - \bar{C})$. Absent this regulation, it would opt for a point such as *Q* that entails a somewhat lower solvency level (compared to \bar{S}_{II}) in return for a substantial saving of costly risk capital (see also Zweifel et al., 2013).

A bank or an insurer with excess solvency, symbolized by the combination (S^+, C^+) , would not need to react to the shock $d\bar{\sigma} > 0$. 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. For the bank or insurer, this causes the locus F of Figure 3 to become flatter, such as F'' . Now $\bar{C}'' < \bar{C}$ suffices to reach the prescribed solvency level, and the ‘critical’ firm that was at \bar{S}_{II} initially can reduce capital by as much as $(\bar{C} - \bar{C}'')$. This of course holds true of $d\bar{\mu} > 0$ as well.

In sum, one has the following set of conditional predictions (in absolute value) for *Basel II/Solvency II* (subscript *II*), focusing on the critical changes $d\bar{\mu} < 0$ and $d\bar{\sigma} > 0$,

$$0_I < \left| \frac{\partial^2 C}{\partial S \partial \bar{\mu}} \right|_{II} < \left| \frac{\partial^2 C}{\partial S \partial \bar{\mu}} \right|_{S^*} \quad \text{if } d\bar{\mu} < 0 \text{ and } S = \bar{S}_{II}$$

$$0_I < \left| \frac{\partial^2 C}{\partial S \partial \bar{\sigma}} \right|_{II} < \left| \frac{\partial^2 C}{\partial S \partial \bar{\sigma}} \right|_{S^*} \quad \text{if } d\bar{\sigma} > 0 \text{ and } S = \bar{S}_{II} \tag{21}$$

Here, 0_I symbolizes the zero restriction imposed by *Basel I/Solvency I* [see eq. (18) again]. Applied to eqs. (16) and (17), respectively and in view of assumptions B8 and B9, these restrictions again cause the numerator to increase and the denominator, to decrease. One therefore obtains,

$$\left. \frac{d\hat{\mu}}{d\hat{\sigma}} \right|_{II} > \left. \frac{d\hat{\mu}}{d\hat{\sigma}} \right|_{S^*} \quad \text{if } d\bar{\sigma} > 0 \text{ or } d\bar{\mu} < 0 \text{ and } S = \bar{S}_{II} \tag{22}$$

Figure 2 illustrates once more. *Basel II/Solvency II* being less stringent (at least as long as the required solvency level is not stepped up), its pertinent efficiency frontier runs higher than that of *Basel I/Solvency I* but still lower than absent regulation. To make up for reduced expected returns, even a strongly risk-averse senior management (preferences of type A) is predicted to opt for a more risky allocation $(\sigma_{II}^* > \sigma_{S^*}^*)$ provided the firm just satisfied the *Basel II/Solvency II* norm initially. This condition presumably holds as a rule for those banks and insurance companies with a less risk-averse management (preferences of type B), again resulting in an investment policy that entails a higher volatility of returns than without regulation as well as a reduced solvency level under rather general conditions (see the argument below Conclusion 2 again). In comparison with *Basel I/Solvency I*, these self-defeating effects are less pronounced, since *Basel II/Solvency II* causes a smaller downward shift of the efficiency frontier (see Figure 2 again). However, note once more that the magnitude of this shift typically differs between banks and insurers because it involves $\partial^2 C / \partial S \partial \bar{\mu}$ and $\partial^2 C / \partial S \partial \bar{\sigma}$, parameters that may well take on different values for banks and insurers. The conclusion reached for *Basel I/Solvency I*, stating that solvency regulation needs to be tailored to the specifics of the two types of firms, therefore holds for *Basel II/Solvency II* as well.

In sum, *Basel/Solvency I* and *II* are predicted to have similar effects in one respect. Both may induce banks and insurance companies to opt for more rather than less risk exposure and lower solvency than if they were optimizing free of the respective restraints. However, the two regulations differ in another respect. *Basel I/Solvency I* causes a ‘deformation’ of the (μ, σ) -frontier for all banks and insurance companies. By way of contrast, the ‘deformation effect’ of *Basel II/Solvency II* is limited to the subsets of them who just satisfied the norm initially.

Conclusion 3: Banks and insurance companies just compliant initially with the solvency norm may be induced by *Basel II (Solvency II, respectively)* to pursue a riskier investment policy than absent regulation, but less so than under *Basel I (Solvency I, respectively)*.

6.3 *Basel III/Solvency III*

The details of implementation of *Basel III* regulation are not known yet at the time of writing (this is even more true of *Solvency III*). However, both seek to prescribe a higher level of solvency, to be attained by more solvency capital, an increased part of which shall be equity [Bank for International Settlements (2011); Basel Committee on Banking Supervision (2011)]. In terms of Figure 3, the mandated solvency level shifts toward S^+ or even beyond. The consequence of this shift is that the set of banks and insurers that do not have to react to a shock $d\bar{\sigma} > 0$ shrinks while the set of those who have to come up with the full additional amount of capital to meet *Basel III/Solvency III* requirements increases (see the discussion in Section 6.2 again). Therefore, the steepening of the efficiency frontier predicted in eq. (19) applies to a larger subset of banks and insurance companies, at least during a (lengthy) time of transition that sees them striving to increase their equity and reserves. Moreover, the fact that an increased share of solvency capital must be equity means an increase in regulatory stringency, causing the endogenous efficiency frontier pertaining to *Basel III/Solvency III* to be shifted back down towards that of *Basel I/Solvency I* (see Figure 2 again). Conclusion 3 therefore is predicted to hold more generally, implying that more banks and insurers may in fact be induced to pursue a riskier investment policy resulting in lowered solvency than in the absence of regulation.

7. Summary and Conclusion

The basic hypothesis of this paper states that banks' investment and insurers' investment and underwriting divisions seek to attain a 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. The basic assumption is that these divisions maximize the risk-adjusted rate of return of capital (*RAROC*), with the marginal benefit of a higher level of solvency being the lower cost of refinancing while its marginal cost consisting of returns forgone on the extra capital to be allocated. The divisions of both banks and insurance companies learn the slope of their efficiency frontier in (μ, σ) -space in the course of three periods. In period 1, changes in the market environment occur, viz. an exogenous change in expected returns and in the volatility of returns. These changes induce adjustments during period 2, predicted by comparative static analysis. In period 3, effective adjustment of solvency in turn acts like an exogenous change inherited from period 2, triggering a reallocation of assets (and risks underwritten in the case of insurers) and hence endogenous adjustments in the firm's expected returns and volatility of returns.

These adjustments, derived from 'reverse comparative statics', define the slope of a perceived endogenous frontier in (μ, σ) -space prior to solvency regulation. However, this slope depends importantly on the fact that the relationship between solvency and risk capital and premium income, respectively, is modified by exogenous changes in expected returns and volatility occurring in the capital market (Conclusion 1). The regulations imposed by *Basel I/Solvency I* are now shown to neglect this dependence on market conditions, causing a modification of the (μ, σ) -frontier as perceived by regulated banks and insurance companies. This modification may well induce their senior managements to take a more risky position than it would absent regulation, likely resulting in a lower solvency level (Conclusion 2). However, the magnitude of this effect depends on parameters that reflect the different business models of banks and insurance companies. The implications of *Basel II/Solvency II* are more complex. Still, banks and insurers initially just attaining the prescribed solvency level are again predicted to react to regulation by taking a more risky position than they would have otherwise, with the magnitude of this effect depending on parameters specific to their industry (Conclusion 3). As to *Basel III/Solvency III*, their likely effect will be to increase the subset of banks and insurers taking on more rather than less risk, generally resulting in a reduced solvency level. This can be said to run counter the very objective of the regulators, who want banks and insurance companies to step up rather than lower their solvency levels.

All of these predicted adjustments may be considered to amount to regulatory failure. However, it would be inappropriate to conclude that *Basel I, II or III (Solvency I, II, or III, respectively)* or even solvency regulation in general should be revoked. First, the model analyzed in this paper might be too simplistic; banks and insurers possibly pursue other objectives than just maximizing *RAROC*. Second, *Basel II/Solvency II* already constitutes an improvement over *Basel I/Solvency I* in that its self-defeating effect is limited to the (usually small) subset of banks and insurance companies that initially had just been compliant with the prescribed solvency level. As to *Basel III/Solvency III*, the short-run increase in this problematic subset effect has to be weighed against the long-run increase in solvency levels generally achieved. 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 expected returns and volatility), one would have to find an alternative whose benefit-cost ratio beats that of the *Basel/Solvency* type [for reform proposals, see e.g. Benink and Benston (2005), pp. 311-321]. However, the present work does call attention to likely shortcomings of current and planned future solvency regulation treating banks and insurance companies in the same way [see Zweifel and Eisen (2012), ch. 8.4].

References

- Basel Committee on Banking Supervision. (1988). International convergence of capital measurement and capital standards. Bank for International Settlements, Basel.
- Basel Committee on Banking Supervision. (2011). Summary table. Retrieved February 23, 2015, from <http://www.bis.org/bcbs/Basel3/b3summarytable.pdf>
- Basel Committee on Banking Supervision. (2004). International convergence of capital measurement and capital standards: A revised framework. Bank for International Settlements, Basel.
- Bauer, W., & Ryser, M. (2004). Risk management strategies for banks. *Journal of Banking and Finance*, 28, 331-352. <http://dx.doi.org/10.1016/j.jbankfin.2002.11.001>
- Brandtner, M. (2013). Conditional value-at-risk, spectral risk measures and (non-) diversification in portfolio selection problems - a comparison with mean-variance analysis. *Journal of Banking and Finance*, 37(12), 5526-5537. <http://dx.doi.org/10.1016/j.jbankfin.2013.02.009>
- Covitz, D.M., Hancock, D., & Kwast, M.L. (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.
- Cummins, J.D. (1988). Risk-based premiums for insurance guaranty funds. *Journal of Finance*, 63(4), 823-839. <http://dx.doi.org/10.1111/j.1540-6261.1988.tb02607.x>
- Cummins, J.D., & Nye, D.J. (1981). Portfolio optimization models for property-liability insurance companies. An analysis and some extensions, *Management Science*, 27, 414-430. <http://dx.doi.org/10.1287/mnsc.27.4.414>
- Cummins, J.D., & Phillips, R.D. (2001). Applications of financial pricing models in property-liability insurance. In Dionne, G. (Ed.), *Handbook of Insurance* (ch. 19). Norwell MA, Kluwer.
- Cummins, J.D., & Sommer, D.W. (1996). Capital and risk in property-liability insurance markets. *Journal of Banking and Finance*, 20, 1069-1092. [http://dx.doi.org/10.1016/0378-4266\(95\)00044-5](http://dx.doi.org/10.1016/0378-4266(95)00044-5)
- Cummins, J.D., & Weiss, M.A. (2011). Systemic risk and the U.S. insurance sector. *Working paper series*, Philadelphia, PA: Temple University.
- Davidson, W.N. III, Cross, M.L., & Thornton, J.H. (1992). Corporate demand for insurance: some empirical and theoretical results. *Journal of Financial Services Research*, 6, 61-72. <http://dx.doi.org/10.1007/BF01046118>
- Diamond, D.W., & Dybvig, P.P. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91, 401-419. <http://dx.doi.org/10.1086/261155>
- Epermanis, K., & Harrington, S.E. (2006). Market discipline in property/casualty insurance: evidence from premium growth surrounding changes in financial strength ratings. *Journal of Money, Credit and Banking*, 38(6), 1515-1544. <http://dx.doi.org/10.1353/mcb.2006.0081>
- Flannery, M.J., & Sorescu, S.M. (1996). Evidence of bank market discipline in subordinated debenture yields: 1983-1991. *Journal of Finance*, 51, 1347-1377. <http://dx.doi.org/10.1111/j.1540-6261.1996.tb04072.x>
- Furfine, C.H. (2003). Interbank exposures: Quantifying the risk of contagion. *Journal of Money, Credit, and Banking*, 35, 111-128. <http://dx.doi.org/10.1353/mcb.2003.0004>
- Goldberg, L.G., & Hudgins, S.C. (1996). Response of uninsured depositors to impending S&L failures: Evidence of depositor discipline. *Quarterly Review of Economics and Finance*, 36, 311-325. [http://dx.doi.org/10.1016/S1062-9769\(96\)90018-6](http://dx.doi.org/10.1016/S1062-9769(96)90018-6)
- Goldberg, L.G., & Hudgins, S.C. (2002). Depositor discipline and changing strategies for regulating thrift institutions. *Journal of Financial Economics*, 63, 263-274. [http://dx.doi.org/10.1016/S0304-405X\(01\)00096-4](http://dx.doi.org/10.1016/S0304-405X(01)00096-4)
- Jacklin, C.J., & Bhattacharya, S. (1988). Distinguishing panics and information-based bank runs: Welfare and policy implications. *Journal of Political Economy*, 96, 568-592. <http://dx.doi.org/10.1086/261552>
- Jensen, M.C., & Meckling, W.H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership

- structure. *Journal of Financial Economics*, 3, 305-360. [http://dx.doi.org/10.1016/0304-405X\(76\)90026-X](http://dx.doi.org/10.1016/0304-405X(76)90026-X)
- John, K., Saunders, A., & Senbet, L.W. (2000). A theory of bank regulation and management compensation. *Review of Financial Studies*, 13(1), 95-125. <http://dx.doi.org/10.1093/rfs/13.1.95>
- Jordan, J.S. (2000, March/April). Depositor discipline at failing banks. *New England Economic Review*, 15-28.
- Kim, D., & Santomero, A. (1988). Risk in banking and capital regulation. *Journal of Finance*, 43(5), 1219-1233. <http://dx.doi.org/10.1111/j.1540-6261.1988.tb03966.x>
- Koehn, M., & Santomero, A. (1980). Regulation of bank capital and portfolio risk. *Journal of Finance*, 35(5), 1235-1244. <http://dx.doi.org/10.1111/j.1540-6261.1980.tb02206.x>
- Lang, L.H.P., & Stulz, R.M. (1992). Contagion and competitive intra-industry effects of bankruptcy announcements. *Journal of Financial Economics*, 32, 45-60. [http://dx.doi.org/10.1016/0304-405X\(92\)90024-R](http://dx.doi.org/10.1016/0304-405X(92)90024-R)
- Merton, R.C. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*, 29, 449-470. <http://dx.doi.org/10.1111/j.1540-6261.1974.tb03058.x>
- Merton, R.C. (1977). An analytical derivation of the cost of deposit insurance and loan guarantees. *Journal of Banking and Finance*, 1, 3-11. [http://dx.doi.org/10.1016/0378-4266\(77\)90015-2](http://dx.doi.org/10.1016/0378-4266(77)90015-2)
- Mülbert, P.O. (2009). Corporate governance of banks. *European Business Organization Law Review*, 10, 411-436. <http://dx.doi.org/10.1017/S156675290900411X>
- Park, S., & Peristiani, S. (1998). Market discipline by thrift depositors. *Journal of Money, Credit, and Banking*, 30, 347-364. <http://dx.doi.org/10.2307/2601105>
- Repullo, R. (2004). Capital requirements, market power, and risk-taking banking. *Journal of Financial Intermediation*, 13, 156-182. <http://dx.doi.org/10.1016/j.jfi.2003.08.005>
- Power, M. (2004). *The Risk Management of Everything. Rethinking the Politics of Uncertainty*. Demos, London.
- Rochet, J.-C., 1992. Capital requirements and the behaviour of commercial banks. *European Economic Review*, 36, 1137-1178.
- Shrieves, R.E., & Dahl, D. (1992). The relationship between risk and capital in commercial banks. *Journal of Banking and Finance*, 16(2), 439-457. [http://dx.doi.org/10.1016/0378-4266\(92\)90024-T](http://dx.doi.org/10.1016/0378-4266(92)90024-T)
- Smith, C.W., & Stulz, R.M. (1985). The determinants of firms' hedging policies. *Journal of Financial and Quantitative Analysis*, 20, 391-405. <http://dx.doi.org/10.2307/2330757>
- Stulz, R.M. (1996). Rethinking risk management. *Journal of Applied Corporate Finance*, 9, 8-24. <http://dx.doi.org/10.1111/j.1745-6622.1996.tb00295.x>
- Woehrmann, P., Semmler, W., & Lettau, M. (2004). Nonparametric estimation of the time-varying Sharpe ratio in dynamic asset pricing models. *Institute for Empirical Research Working Paper No. 225*, University of Zurich, Zurich.
- Zweifel, P. (2014). Solvency regulation of insurers: A regulatory failure?. *Journal of Insurance Issues*, 37(2), 135-157.
- Zweifel, P., & Eisen, R. (2012). *Insurance Economics*. Springer, Heidelberg. <http://dx.doi.org/10.1007/978-3-642-20548-4>
- Zweifel, P., Pfaff, D., & Kühn, J. (2013). A simple model of bank behavior – with implications for solvency regulation. *Studies in Microeconomics*, 3(1).

Appendix A. Bank model

First, consider a shock $d\bar{\mu}$ disturbing the first-order condition (4). With R shorthand for $RAROC$, the comparative static equation reads,

$$\frac{\partial^2 R}{\partial S^2} dS^* + \frac{\partial^2 R}{\partial S \partial \bar{\mu}} d\bar{\mu} = 0. \tag{A.1}$$

Since $\partial^2 R / \partial S^2 < 0$ in the neighborhood of a maximum, $\text{sgn}[\partial^2 R / \partial S \partial \bar{\mu}]$ determines $\text{sgn}[dS^* / d\bar{\mu}]$. Differentiating eq. (4) w.r.t. $\bar{\mu}$, one has

$$\frac{\partial^2 R}{\partial S \partial \bar{\mu}} = -\frac{\partial^2 r_D}{\partial S \partial \bar{\mu}} - \left(\left(1 - \frac{\partial r_D}{\partial \bar{\mu}} \right) C - (\mu - r_D) \frac{\partial C}{\partial \bar{\mu}} \right) \frac{1}{C^2} \cdot \frac{\partial C}{\partial S} - \frac{\mu - r_D}{C} \cdot \frac{\partial^2 C}{\partial S \partial \bar{\mu}}. \tag{A.2}$$

Using (4) to obtain $\partial C / \partial S = -(\mu - r_D)^{-1} (\partial r_D / \partial S) \cdot C$, one has

$$\frac{\partial^2 R}{\partial S \partial \bar{\mu}} = -\frac{\partial^2 r_D}{\partial S \partial \bar{\mu}} + \left(1 - \frac{\partial r_D}{\partial \bar{\mu}} \right) \frac{1}{C} \cdot \frac{1}{\mu - r_D} \cdot C \cdot \frac{\partial r_D}{\partial S} - (\mu - r_D) \frac{\partial C}{\partial \bar{\mu}} \cdot \frac{1}{C^2} \cdot \frac{1}{\mu - r_D} \cdot C \cdot \frac{\partial r_D}{\partial S} - \frac{\mu - r_D}{C} \cdot \frac{\partial^2 C}{\partial S \partial \bar{\mu}}. \tag{A.3}$$

This can be simplified to become eq. (9) of the text.

Now consider $d\bar{\sigma} > 0$. In full analogy to (A.1), one obtains from eq. (4),

$$\left[\frac{\partial^2 R}{\partial S \partial \bar{\sigma}} \right] = -\frac{\partial^2 r_D}{\partial S \partial \bar{\sigma}} - \left(-\frac{\partial r_D}{\partial \bar{\sigma}} \cdot C - (\mu - r_D) \cdot \frac{\partial C}{\partial \bar{\sigma}} \right) \cdot \frac{1}{C^2} \cdot \frac{\partial C}{\partial S} - \frac{\mu - r_D}{C} \cdot \frac{\partial^2 C}{\partial S \partial \bar{\sigma}}. \tag{A.4}$$

Using (4) again to substitute $\partial C / \partial S$, one has

$$\frac{\partial^2 R}{\partial S \partial \bar{\sigma}} = -\frac{\partial^2 r_D}{\partial S \partial \bar{\sigma}} - \frac{\partial r_D}{\partial \bar{\sigma}} \cdot \frac{1}{C} \cdot \frac{1}{\mu - r_D} \cdot C \cdot \frac{\partial r_D}{\partial S} - (\mu - r_D) \frac{\partial C}{\partial \bar{\sigma}} \cdot \frac{1}{C^2} \cdot \frac{1}{\mu - r_D} \cdot C \cdot \frac{\partial r_D}{\partial S} - \frac{\mu - r_D}{C} \cdot \frac{\partial^2 C}{\partial S \partial \bar{\sigma}}. \tag{A.5}$$

Slight rearrangement yields eq. (11) of the text.

Appendix B. Insurer model

First, consider a shock $d\bar{\mu}$ disturbing the first-order condition (4). As stated in

Appendix A, $\text{sgn}[\partial^2 R / \partial S \partial \bar{\mu}]$ determines $\text{sgn}[dS^* / d\bar{\mu}]$. Differentiating eq. (7) w.r.t. $\bar{\mu}$, one has

$$\begin{aligned} \frac{\partial^2 R}{\partial S \partial \bar{\mu}} &= \frac{\partial e(P, S)}{\partial \bar{\mu}} - \frac{\partial e(C, S)}{\partial \bar{\mu}} = \frac{\partial}{\partial \bar{\mu}} \left[\frac{\partial P}{\partial S} \cdot \frac{S}{P} \right] - \frac{\partial}{\partial \bar{\mu}} \left[\frac{\partial C}{\partial S} \cdot \frac{S}{C} \right] \\ &= \frac{\partial^2 P}{\partial S \partial \bar{\mu}} \cdot \frac{S}{P} + \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial P / \partial \bar{\mu}}{P^2} \right) - \frac{\partial^2 C}{\partial S \partial \bar{\mu}} \cdot \frac{S}{C} - \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial C / \partial \bar{\mu}}{C^2} \right) \end{aligned} \tag{B.1}$$

After slight rearrangement, this is expression (13) of the text.

Now consider $d\bar{\sigma} > 0$. In full analogy to (B.1), one obtains from eq. (7),

$$\begin{aligned} \frac{\partial^2 R}{\partial S \partial \bar{\sigma}} &= \frac{\partial e(P, S)}{\partial \bar{\sigma}} - \frac{\partial e(C, S)}{\partial \bar{\sigma}} = \frac{\partial}{\partial \bar{\sigma}} \left[\frac{\partial P}{\partial S} \cdot \frac{S}{P} \right] - \frac{\partial}{\partial \bar{\sigma}} \left[\frac{\partial C}{\partial S} \cdot \frac{S}{C} \right] \\ &= \frac{\partial^2 P}{\partial S \partial \bar{\sigma}} \cdot \frac{S}{P} + \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial P / \partial \bar{\sigma}}{P^2} \right) - \frac{\partial^2 C}{\partial S \partial \bar{\sigma}} \cdot \frac{S}{C} - \frac{\partial P}{\partial S} \left(\frac{-S \cdot \partial C / \partial \bar{\sigma}}{C^2} \right). \end{aligned} \tag{B.2}$$

After slight rearrangement, this is expression (15) of the text.