Rebalancing the 3 Pillars of Basel 2∗

Jean-Charles Rochet†

September 26, 2003

Abstract: This paper develops a formal model of banking regulation where the interactions between the 3 pillars of Basel 2 can be analyzed. It also provides a formal analysis of the mandatory subdebt proposal and shows that, under certain conditions, Pillar 1 (solvency ratio) and Pillar 3 (market discipline) are partially substitutable. The same is true for Pillar 2 (supervisory action) and Pillar 3. Moreover, I find a strong complementarity between market discipline and supervisory action: one cannot work efficiently without the other. The main policy implication is that, instead of spending so much time on refining Pillar 1, the Basel Committee should seriously think about the practical ways to implement Pillars 2 and 3.


†Toulouse University and L.S.E., Email: rochet@cict.fr.
1 Introduction

The on-going reform of the Basel Accord is supposed to rely on three “pillars”: a new capital ratio, supervisory review and market discipline. But even a cursory look at the proposals of the Basel Committee on Banking Supervision (BCBS) reveals a certain degree of imbalance between these three pillars. Indeed, the BCBS gives a lot of attention to the refinements of the risk weights in the new capital ratio (132 pages in the 3rd Consultative Paper of April 2003) but is much less precise about the other pillars (16 pages on Pillar 2 and 15 pages on Pillar 3).\(^1\)

Even though the initial capital ratio (Basel Committee, 1988) has been severely criticized for being too crude and opening the door to regulatory arbitrage,\(^2\) it seems strange to insist so much on the importance of supervisory review\(^3\) and market discipline as necessary complements to capital requirements, while remaining silent of the precise ways\(^4\) this complimentarity can work in practice. One possible reason for this imbalance is a gap in the theoretical literature. As far as I know, there is no tractable model that allows a simultaneous analysis of the impact of solvency regulations, supervisory action and market discipline on commercial banks’ behavior.

This article aims at filling this gap by providing a simple analytical framework for analyzing the interactions between the three pillars of Basel 2. I start by a critical assessment of the academic literature\(^5\) on the three pillars (Section 2). I argue that none of the existing models allows for a satisfactory integration of these 3 pillars. I therefore develop a new formal model (Section 3) that tries to incorporate the most important criticisms that have been addressed to existing theoretical models of bank regulation. In Section 4, I show that minimum capital ratios can be justified by a classical agency problem à la Holmström (1979) between bankers and regulators, even in the absence of mispriced deposit insurance. In Section 5, I show that, under restrictive conditions, these capital requirements can be reduced if banks are mandated to issue subordinated debt on a regular basis (direct market discipline). In Section 6, I explore the interactions between market discipline and supervisory action and show that they are complements rather than

---

\(^1\)This imbalance is also reflected in the comments on Basel 2: see Sainenberg and Schuermann (2003) for an assessment.


\(^3\)For example, the BCBS insists on the need to “enable early supervisory intervention if capital does not provide a sufficient buffer against risk” (Basel Committee, 2003).

\(^4\)In particular, in spite of the existence of very precise proposals by US economists (Calomiris (1998), Evanoff and Wall (2000), see also the discussion in Bliss (2001)) for mandatory subordinated debt, these proposals are not discussed in the Basel 2 project.

\(^5\)Good reviews of this literature can be found in Thakor (1996), Jackson et al. (1999) and Santos (2000).
2 The 3 Pillars in the Academic Literature

Most of the academic literature that commented on Basel 1 has concentrated on the “credit crunch” of the early 1990s\(^6\) and on the distortions of banks’ asset allocation generated by the wedge between the market assessment of banks’ asset risks and its regulatory counterpart in Basel I. Several theoretical articles (e.g., Koehn and Santomero (1980), Kim and Santomero (1988), Furlong and Keeley (1990), Rochet (1992) and Thakor (1996)) used static portfolio models to explain these distortions. In such models, appropriately designed capital requirements could be used to correct the incentives of bank shareholders to take excessive risks, due to the mispricing of deposit insurance of simply to the limited liability option. Using a different approach, Froot and Stein (1998) model the buffer role of bank capital in absorbing liquidity risks. They determine the capital structure that maximizes the bank’s value when there are no audits nor deposit insurance.

Yet, as pointed out by Hellwig (1998) and Blum (1999), static models fail to capture important effects of capital requirements. The empirical literature (e.g., Hancock et al. (1995) and Furfine (2001) has tried to calibrate dynamic models of bank behavior, in order to study those intertemporal effects. However, none of these papers study the interactions of capital requirements with supervisory action\(^7\) and market discipline.

The early literature on continuous time models of banks’ behavior was initiated by Merton (1977). Assuming an exogenous closure date, Merton (1977) shows that the fair price of deposit insurance can be computed as a European put option. Merton (1978) extends this framework by considering random audits and endogenous closure dates.\(^8\)

Merton’s seminal contributions have extended in several directions:

- Fries et al. (1997) introduce deposit withdrawal risk and study the impact of banks closure policy on the fair pricing of deposit insurance.

- Bhattacharya et al. (2002) study closure rules that can be contingent on the level of risk taken by the bank.

---

\(^6\)Good discussions of this issue can be found in Berger and Udell (1994), Thakor (1996), Jackson et al. (1999) and Santos (2000).

\(^7\)However, Peek and Rosengren (1995) provide empirical evidence for the impact of increased supervision on bank lending decisions.

\(^8\)However, Merton (1977, 1978) assumes that banks’ assets are traded on financial markets (in order to use the arbitrage pricing methodology), which implies that the social value of banks is independent of their liability structure.
Levonian (2001) introduces subordinated debt in Merton (1977)'s model and studies its impact on bankers' incentives for risk taking.

3 A Formal Model

Our model\(^9\) aims at taking seriously most of the criticisms addressed to previous models of bank regulation, while remaining tractable.

First, it is a dynamic model, because static models necessarily miss important consequences of bank solvency regulations.\(^{10}\) The simplest dynamic models are in discrete time, like Calem and Rob (1996) or Buchinsky and Yosha (1997), but they typically don’t yield closed form solutions and entail the use of numerical simulations. For transparency reasons we use instead a continuous time model à la Merton (1977, 1978), which requires using diffusion calculus but ultimately reveals more tractable. Following Merton, we therefore assume that the (economic) value \(A_t\) of a bank’s assets at date \(t\) follows a diffusion process:

\[
\frac{dA}{A} = \mu dt + \sigma dW, \tag{1}
\]

where \(dW\) is the increment of a Wiener process. \(\mu\) and \(\sigma\) are the drift and volatility of asset value. For simplicity, all investors are risk neutral\(^{11}\) and discount the future at a constant rate \(r > \mu\). The bank’s assets continuously generate an instantaneous cash-flow \(x_t = \beta A_t\), where \(\beta > 0\) is the constant pay-out rate.\(^{12}\)

We depart from the complete frictionless markets assumption made by Merton (1977, 1978)\(^{13}\) and many of his followers, since this assumption implies that the social value of banks is independent of their liability structure, a hardly acceptable feature if one wants to study the consequences of solvency regulations for banks. Following Gennotte and Pyle (1991), we assume instead that banks create value by monitoring borrowers, and thus acquire private information about these borrowers. The counterpart of this private

---

\(^9\)It is a variant of the model used by Décamps et al. (2003) who also analyze the interactions between the 3 pillars of Basel 2.

\(^{10}\)The main problem is that, in a static model, solvency regulations only have an impact when they are binding. However the great majority of banks have today (this was not the case in 1988) much more capital than the regulatory minimum. Hence it is difficult to explain the impact of a capital ratio in today’s world by using a static model. For other critiques, see Blum (1999).

\(^{11}\)Small depositors are risk averse but they are fully insured and do not play any active role in our model. In any case, risk neutrality is not crucial for our results but it simplifies the analysis. We could assume alternatively that all investors use the same risk adjusted measure for evaluating risky cash flows.

\(^{12}\)Because of risk neutrality, the expected net present value of these cash flows (conditional on the information available at date \(t\)) has to coincide with \(A_t\), which is easily seen to be equivalent to the condition \(\beta = r - \mu\).

\(^{13}\)In Merton (1978) there are audit costs but no liquidation costs due to the resale of banks’ assets, like we have here. As a result, social surplus is unaffected by liquidation decisions.
information is that the resale value of a bank’s assets (typically, in case of liquidation) is only a fraction $\lambda A$ (with $\lambda < 1$) of the economic value $A$ of these assets.\(^{14}\)

We also assume that monitoring borrowers has a fixed cost component,\(^{15}\) equivalent to a flow cost $r\gamma$ per unit of time (its present value is thus $\gamma$). From a social perspective a bank should thus be closed when its asset value falls below some threshold $A_L$ (the liquidation threshold). The social value of the bank (denoted $V(A)$) equals the expected present value of future cash flows $x_t = \beta A_t$, net of monitoring costs $r\gamma$, until the stopping time $\tau_L$ (the first time $t$ where the bank hits the liquidation threshold $A_L$) where the bank is liquidated and its assets resold at price $\lambda A_L$. As shown in the Appendix, this social value equals:

$$V(A) = (A - \gamma) + \{\gamma - (1 - \lambda)A_L\} \left(\frac{A}{A_L}\right)^{-a},$$

(2)

where $a$ is the positive root of the quadratic equation:

$$\frac{1}{2}\sigma^2 x^2 - \mu x = r.$$

(3)

Notice that this total value is composed of two terms:

- the value $A$ of its assets, net of monitoring costs $\gamma$;
- the option value associated to the (irreversible) closure decision.

As in the real options literature (see for example Dixit and Pindyck (1994)) the social value of the bank is maximized for a threshold $A_L$ that is below the break-even threshold $A_0 = \frac{\gamma}{1 - \lambda}$. More specifically, the level of $A_L$ that maximizes the social value of the bank (notice that this level is independent of $A$, and thus is time-consistent) is what we call the first-best closure threshold:

$$A_{FB} = \frac{a}{(a + 1)(1 - \lambda)} \frac{\gamma}{1 - \lambda}.$$

(4)

Thus our model captures an important feature of real life banking systems.\(^{16}\) even in the absence of moral hazard, government subsidies and the like, the failure rate of

\(^{14}\)A similar assumption is made in the corporate finance literature (Leland (1994), Leland and Toft (1996), Mella-Barral and Perraudin (1997)), but $(1 - \lambda)$ is interpreted as a “physical” liquidation cost. Here it is an asymmetric information cost, related to the opacity of banks’ assets.

\(^{15}\)A proportional cost of monitoring (if any) can be substracted with the drift in Equation (1).

\(^{16}\)Of course, our desire to get closed form solutions limits our model to one state variable only. This means that we cannot address important questions like the way banks allocate their assets among different classes of risks and the hoarding of liquid assets as another buffer against risk. The first topic is addressed by the vast literature on risk-weighted ratios (Koehn and Santomero (1980), Kim and Santomero (1988), Rochet (1992), Furlong and Keeley (1996) and Thakor (1996)). The second topic is addressed in Milne and Whalley (2001).
banks cannot be zero. The socially optimal failure rate takes into account the embedded real option: given that bank closures are irreversible (and entail a real cost, due to the imperfect resaleability of banks’ assets) it is optimal to let banks operate (up to a certain point) below the break-even level (defined by the condition $A_0 - \gamma = \lambda A_0$) in the hope that they recover. In more concrete terms, the fact that resolution of banks failures costs money (ex post) to the Deposit Insurance Fund does not necessarily imply some kind of inefficiency. This feature is illustrated in Figure 1 below:\footnote{Notice that the first best social value of the bank is a convex function of assets value. However when agency problems are taken into account, and make the liquidation threshold become greater than $A_0$, the value function becomes concave, and thus exhibits risk aversion.}

So far, we have only introduced one of the two important features of banking: banks’ assets are opaque,\footnote{Morgan (2002) provides indirect empirical evidence on this opacity by comparing the frequency of disagreements among bond rating agencies about the values of firms across sectors of activity. He shows that these disagreements are much more frequent, all else being equal, for banks and insurance companies than for other sectors of the economy.} which implies that they have to be monitored and cannot be resold at full value. We now introduce the second feature: the bulk of a bank’s liabilities consists of retail deposits $D$, fully insured by a Deposit Insurance Fund (DIF in the sequel) and paying interest $rD$. The DIF is financed by a premium $P$ paid at discrete dates. We assume that this premium is fair (so we rule out systematic subsidies from the DIF to the banks) but cannot be revised by continuous readjustments. The academic literature has insisted a lot on the “moral hazard” problem created by the put-option feature of deposit insurance. It has been extensively argued that this feature, and more generally the limited

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{The real option embedded in a bank’s social value}
\end{figure}
liability of bank shareholders, gave these shareholders incentives to take excessive risks, especially when banks are insufficiently capitalized. We focus here on a different agency problem, also created by limited liability, but of a different nature: bankers\textsuperscript{19} may not have enough incentives to monitor their assets when the value of these becomes too small.\textsuperscript{20} We assume that when monitoring stops (we say that the banker shirks), the quality of banks’ assets deteriorates,\textsuperscript{21} and the dynamics of asset value becomes:

\[
\frac{dA}{A} = (\mu - \Delta \mu)dt + \sigma dW, \tag{5}
\]

where $\Delta \mu > 0$. Equation (5) means that the shirking/monitoring decision only impacts the expected profitability of the bank’s assets and not their risk. Most of the academic literature has considered the polar case where $\mu$ is unchanged but $\sigma$ increased by moral hazard (assets substitution problem). Which specification is more appropriate is an empirical question.\textsuperscript{22} In Décamps et al. (2003) we consider the general case where both $\mu$ and $\sigma$ are altered by the banker’s decision.

In the absence of shirking, the value of the bank’s equity is given by a simple formula in the spirit of Formula (2) (see the Appendix for all mathematical derivations):

\[
E(A) = A - D - \gamma + (D + \gamma - A_L) \left( \frac{A}{A_L} \right)^{-a}. \tag{6}
\]

As in Merton (1977, 1978) the value of the bank’s equity is the sum of two terms:

- the value of assets $A$ net of debt (deposits) $D$ (and of the monitoring cost, that does not appear in Merton 1974),
- the value of the limited liability option.

However, like in Merton (1978) but in contrast with Merton (1977), the value of the limited liability option is not of the Black-Scholes type (it is actually simpler). This is because it can be exercised at any time (down and out barrier option) instead of at a fixed date (European option). In Merton (1978) the closure option can only be exercised after

\textsuperscript{19}We assume that bank managers act in the best interest of shareholders. It would be interesting, but presumably difficult, to introduce also an agency problem between managers and shareholders.
\textsuperscript{20}Bliss (2001) argues that this agency problem may be fundamental: “Poor (apparently irrational) investments are as problematic as excessively risky projects (with positive risk-adjusted returns). Evidence suggests that poor investments are likely to be the major explanation for banks getting into trouble”.
\textsuperscript{21}We assume that this deterioration, i.e. $\Delta \mu$ in equation (5), is large enough so that shirking is never optimal from a social viewpoint.
\textsuperscript{22}In the appendix to his paper, Bliss (2001) convincingly argues that the assets substitution problem might have been over emphasized. He reviews several empirical articles that conclude that bank failures are often provoked by “bad investments” rather than “bad luck” (and excessive risk taking).
an audit by the supervisor. Here we assume that, thanks to the information revealed by the market (indirect market discipline) closure can occur at any time.

$A_L$ is now chosen by shareholders so as to maximize equity value. The corresponding threshold is:

$$A_E = \frac{a}{a + 1}(D + \gamma).$$  \hfill (7)

At this threshold the marginal value of equity is zero: $E'(A_E) = 0$ (smooth pasting condition).

Shirking is optimal for bankers whenever their expected instantaneous loss from shirking, $AE'(A)\Delta \mu$ becomes less than the instantaneous monitoring cost $\gamma r$. Since $E'(A_E) = 0$, this has to be true on some interval $[A_E, A_S]$.

**Proposition 1**: When the cost of monitoring $\frac{\gamma D}{D}$ (per unit of deposits) is smaller than $\frac{a + 1}{a} - 1$, there is a conflict of interest between shareholders and the FDIC: insufficiently capitalized banks shirk.

Interestingly, there are parameter values for which the agency problem does not matter: when $\frac{\gamma D}{D}$ is large ($> \frac{a + 1}{a} - 1$) shareholders decide to close the bank before the shirking constraint becomes binding. However when $\frac{\gamma D}{D}$ is smaller than this threshold, there is a conflict of interest, even in the absence of mispriced deposit insurance. It is similar to the conflict of interest between bondholders (who typically don’t subsidize firms!) and shareholders of undercapitalized firms.

Figure 2 represents the typical pattern of the value of the bank’s equity $E$, as a function of the value $A$ of its assets, in the case where deposits are fully insured (and therefore depositors have no incentives to withdraw), but the bank is left unregulated. The closure threshold $A_E$ (below which the bank declares bankruptcy)\(^{23}\) and the shirking threshold $A_S$ (below which the bank shirks) are chosen by bankers so as to maximize the value of equity. The reason why shirking is sometimes preferred by shareholders (in the intermediate region $A_E \leq A \leq A_S$) even though it is socially inefficient, is not the deposit insurance option (as is the case for assets substitution problems, where typically $\gamma = \Delta \mu = 0, \Delta \sigma^2 > 0$) but rather the agency problem between the bank and the Deposit Insurance System. Indeed, the cost of monitoring is entirely borne by the bank, but the bank only collects a fraction of the benefits of monitoring. When these benefits are small, the banker prefers to shirk. As we see in the next section, this can be prevented by imposing a minimum capital ratio.

\(^{23}\)Following Leland (1994) we assume that shareholders are not cash constrained. In Décamps et al. (2003) we examine the alternative case where bankruptcy is precipitated by a bank’s liquidity problems.
4 Justifying the Minimum Capital Ratio

In our model, the justification of a minimum capital ratio is not an asset substitution problem (like in the vast majority of academic papers on the topic\textsuperscript{24}) but an agency problem between the banker and the supervisors, who represent the interests of the DIF: insufficiently capitalized banks “shirk”, i.e. stop monitoring their assets. To avoid this problem, we assume now that when the value of the bank’s assets hits some threshold $A_R$ chosen by the regulator, the bank is liquidated and the shareholders are expropriated.\textsuperscript{25} This regulatory threshold $A_R$ is designed in such a way that bankers are never tempted to shirk.

**Proposition 2 : 1)** Under the assumptions of Proposition 1 (i.e., $\frac{\gamma}{D} < \frac{a+1}{\lambda a} - 1$), bank shirking can be eliminated if bank regulators impose liquidation whenever the bank’s assets fall below the following threshold:

$$A_R^Q = \frac{a(D + \gamma) + \frac{\gamma r}{\Delta \mu}}{a + 1}.$$ \hspace{1cm} (8)

**2)** When $\frac{\gamma}{D}$ is larger than $\frac{\Delta \mu}{r + a \Delta \mu}$, this liquidation threshold can be implemented by imposing

\textsuperscript{24}See Santos (2000) for a review.

\textsuperscript{25}This is similar to protected debt covenants, whereby bondholders (or banks) retain the option of restructing a firm before it is technically bankrupt. See Black and Cox (1976) for a formal analysis.
a minimum capital ratio:
\[
\frac{A - D}{A} \geq \rho_R \equiv \frac{\gamma(r + a\Delta\mu) - D\Delta\mu}{\gamma(r + a\Delta\mu) + aD\Delta\mu}.
\] (9)

The condition \( \frac{\gamma}{D} > \frac{\Delta\mu}{r + a\Delta\mu} \) ensures that the minimum capital ratio \( \rho_R \) is positive. When it is not satisfied, banks can be allowed to continue for negative equity values. We focus here on the more interesting case of Proposition 2.2, where \( \rho_R > 0 \).

Notice that when liquidation costs are large enough the first-best liquidation threshold \( A_{FB} = \frac{a}{a+1} \frac{\gamma}{1-\lambda} \) is smaller than the regulatory threshold \( A_{OR} = \frac{a(D+\gamma + \frac{\gamma}{\Delta\mu})}{a+1} \). This means that bank supervisors are confronted with time consistency problems: even if ex-ante, agency considerations imply that a bank should be closed whenever \( A \leq A_{OR} \), ex post it is optimal to let it continue (and provide liquidity assistance). These forbearance problems are examined in Section 6.

To conclude this section, let us examine the policy implications of our first results. We interpret bank solvency regulations as a closure rule intended to avoid shirking by insufficiently capitalized banks: every time the assets value of the banks falls below \( A_{OR} \), the bank should be closed. However, we have argued that banks assets are opaque, and cannot be marked to market in continuous time. The traditional view on supervisors’ role was to evaluate these assets periodically through on site examinations. In particular, most academic papers\(^{26}\) assumed that \( A_t \) was only observable through costly auditing, that had to be performed more or less with uniform frequency across banks. We argue in favor of a more modern approach whereby bank supervisors can rely on market information and adapt the intensity or frequency of their examinations to the market assessment of the bank’s situation. This can be done in several ways: conditioning risk weights on market ratings of assets (Pillar 1), or using yield spreads of bank liabilities and private ratings for reassessing the solvency of the bank (Pillars 2 and 3). In our model, it would mean inferring \( A_t \), the (unobservable) value of the bank’s assets, from the market price of equity (if the bank is publicly listed), that is inverting the function \( A \rightarrow E(A) \). This is the main content of what is called “indirect market discipline”: using as much publicly available information as possible to allocate scarce regulatory resources in priority to the banks in distress. In our model, it translates into a simple policy recommendation of organizing a regulatory framework with two regimes:

- A light regime for “healthy” banks (those for which assets value \( A \) is way above the closure threshold \( A_{OR} \), where \( A \) is inferred from accounting data and market information), only imposing accurate reporting and transparency.

\(^{26}\) The seminal paper on this is Merton (1978). More recent references are Fries et al. (1997) and Bhattacharya et al. (2002).
• A heavy regime for “problem” banks (those for which \( A \) gets close to \( A_O \)) imposing restrictions on what the bank can do, and closely examining its books.

This two-regime regulatory framework (inspired of the Prompt Corrective Action provisions of FDICIA)\(^{27}\) is examined formally in Section 6. For the moment, we discuss direct market discipline, and more specifically how the subdebt proposal can, under certain conditions, reduce capital requirements.

5 Market Discipline and Subordinated Debt

We consider now that the bank is mandated to issue a certain volume \( B \) of bonds,\(^{28}\) each paying a continuous coupon \( c \) per unit of time. These bonds are subordinated to deposits: if the bank is liquidated (when assets value hits threshold \( A_L \)), the DIF receives \( \lambda_A L \) but bondholders receive nothing. Anticipating on this possibility, bondholders require a coupon rate \( c \) above the riskless rate \( r \). To maintain the convenience of the stationarity of the bank’s financial structure, we assume that bonds have an infinite maturity (unless of course the bank is liquidated) but are randomly renewed according to a Poisson process\(^{29}\) of intensity \( m \). In more intuitive terms, a fraction \( m dt \) of outstanding bonds is repaid at each instant at its face value \( B \) and the same volume \( m dt \) is reissued,\(^{30}\) but at its market value \( B(A) \). This is where market discipline comes in.\(^{31}\) if the bank’s asset value \( A \) deteriorates (for example if bankers stop monitoring their assets) the finance cost of the bank increases immediately, since at each instant a fraction \( m \) of the bonds has to be repaid at face value \( B \) and reissued at market value \( B(A) < B \) (notice also that \( B' > 0 \)). In the Appendix we show that the value of equity becomes:

\[
E(A, B) = A - \gamma - D - B \frac{c}{r} + (D + \gamma - A_L) \left( \frac{A}{A_L} \right)^{-a} + \frac{c}{r} B \left( \frac{A}{A_L} \right)^{-a(m)},
\]

(10)

where \( a(m) \) is the positive root of the quadratic equation

\[
\frac{1}{2} \rho^2 x^2 - \mu x = r + m.
\]

\(^{27}\) The consequences of FDICIA are assessed in Jones and King (1995) and Mishkin (1996).

\(^{28}\) The mandatory subdebt proposal has been discussed extensively: see e.g. Calomiris (1998), Estrella (2000), Evanoff and Wall (2000, 2001). The only formal analysis I am aware of is Levonian (2001). However Levonian (2001) uses a Black-Scholes type of model where the bank’s returns on assets are exogenous. For empirical assessments of the feasibility of the subdebt proposal, see Hancock and Kwast (1996) and Sironi (2001).

\(^{29}\) This trick is borrowed from Leland and Toft (1996) and Ericsson (2000).

\(^{30}\) The average maturity of bonds is thus \( \int_0^\infty m e^{-mt} dt = \frac{1}{m} \).

\(^{31}\) The disciplining role of periodical repricing of debt has been shown be Levonian (2001). It is close in spirit to the disciplining role of demandable deposits (Calomiris and Kahn, (1991), Carletti (1999)).
The new regulatory threshold $A^{SD}_R$ is the smallest value of $A_L$ that guarantees that bankers will not shirk. It is defined implicitly by:

$$A^{SD}_R \frac{\partial E}{\partial A}(A^{SD}_R, B) = \frac{\gamma}{\Delta \mu},$$

where $A_L$ is taken to be equal to $A^{SD}_R$. After easy computations we obtain:

$$A^{SD}_R = A^O_R + \frac{a(m)}{a+1} c r B. \quad (11)$$

Not surprisingly, this threshold is higher than $A^O_R$ (all else being fixed), since the bank is now more indebted. However the capital ratio becomes:

$$\rho^{SD}_R = 1 - \frac{D + B}{A^{SD}_R}. \quad (12)$$

It is smaller than $\rho_R = 1 - \frac{D}{A^O_R}$, whenever

$$\frac{A^{SD}_R}{D + B} < \frac{A^O_R}{D},$$

which is equivalent to:

$$a(m) \frac{c}{r} < a + \frac{\gamma}{D} \left( a + \frac{r}{\Delta \mu} \right). \quad (13)$$

Since $a(m)$ increases with $m$ (which $a(0) = 0$ and $a(+\infty) = +\infty$), we see that this condition can be satisfied, but $m$ (the frequency of renewal of bonds, which is inversely proportional in our model to their average maturity) and $\frac{c}{r}$ (the relative spread on subordinated debt) have to be small enough. Thus we obtain:

**Proposition 3**: If banks are mandated to issue subordinated bonds on a regular basis, regulators can reduce capital requirements (TIER 1) if two conditions are satisfied: the average maturity must not be too small (small) and the coupon paid on the bonds must not be too large ($\frac{c}{r}$ close to 1).

However, the total requirement capital + subdebt (TIER 1 + TIER 2) is always increased.

Proposition 3 shows the limits of the mandatory subdebt proposal. Only when properly designed (i.e. with a not too small maturity or a not too large frequency of renewal) and when markets are sufficiently liquid and bank assets not too risky (so that the relative spread $\frac{c}{r}$ is not too large) can mandatory subdebt allow regulators to decrease capital requirements.

---

32 Recall that in our model, bonds have theoretically an infinite maturity but are randomly repaid with frequency $m$. The average (effective) maturity is $1/m$. 

12


6 Market Discipline and Supervision Action

We come now to what I consider the most convincing rationale for market discipline: preventing regulatory forbearance by forcing regulators to intervene before it is too late. As has been well documented in the literature on banking crises, banking authorities are very often subject to political pressure for bailing out the creditors of banks in distress. To capture this in our model, consider what would happen if subdebt holders where de facto insured in the case where the bank is liquidated (i.e. whenever the value of its assets hits the regulatory threshold). This bail-out obviously does not affect the social value of the bank but only leads to a redistribution of wealth between the DIF and bondholders. Bonds become riskless and perfect substitutes to deposits from the point of view of equity holders. The frequency of renewal of bonds ceases to play any disciplining role.

Proposition 4: If subdebt holders are insured by the DIF, subdebt ceases to play any disciplining role: Direct market discipline can only work in the absence of regulatory forbearance.

Proposition 4 shows the existence of some form of complementarity between market discipline and supervisory action: direct market discipline can only work if supervisors can credibly commit not to bail out bondholders. We now examine a second form of complementarity between market discipline and supervisory action: if financial markets are sufficiently efficient and liquid, and if banks issue publicly traded securities (equity or bonds) the market prices (or yields) of these securities provide objective signals about the situation of these banks. We will not enter the difficult statistical question of which security (equity or subordinated bonds) gives the most useful information to bank supervisors. Our model only has one state variable $A$, and both the equity price $E(A)$ and the subdebt price $B(A)$ are monotonic functions of $A$ and thus sufficient statistics for $A$. In our simple model, by observing market prices of either equity or bonds, regulatory authorities can perfectly infer the true value of $A$.

Our model is obviously not appropriate for analyzing such statistical considerations. It is however well adapted to study another, equally important question, namely how market discipline can limit forbearance. Market information is then viewed as providing objective signals that oblige supervisors to intervene. Indirect market discipline is thus useful in two ways: it allows supervisors to save on audit costs for the banks that are well capitalized, and conversely it forces supervisors to intervene early enough when a bank is in trouble. This is captured in our model in the following way.

---

33 Arbitrage considerations then imply that the coupon rate $c$ must then equal $r$.

34 On this, see for example Bliss (2001), Evanoff and Wall (2002), Gropp et al. (2002) and the references therein.
We consider that bank supervisors are required to inspect banks whenever their assets’ value hit an inspection threshold $A_I$ (with $A_I > A_R$). Inspection allows them to detect shirking, and close the banks who do shirk. The value of equity is still given by formula (6) (for simplicity we go back to the case with no subordinated debt):

$$E(A) = A - \gamma - D + (D + \gamma - A_R) \left( \frac{A}{A_R} \right)^{-a}.$$  

But now the condition for no shirking becomes

$$\forall A \geq A_I, \quad AE'(A) \geq \frac{\gamma r}{\Delta \mu},$$  

which is equivalent to:

$$A_I - a(D + \gamma - A_R) \left( \frac{A_I}{A_R} \right)^{-a} \geq \frac{\gamma r}{\Delta \mu}.$$  

This formula shows that for any closure threshold $A_R$ there is a minimum inspection threshold $A_I$ that prevents shirking: it is given by equality in formula (15). The corresponding curve in the $(A_R, A_I)$ plane is represented in Figure 3. Notice that the previous case (no inspection, closure at $A_R^O$) corresponds to the intersection of this curve with the first diagonal ($A_R = A_I = A_R^O$).

---

**Figure 3:** The optimal combination of inspections (threshold $A_I^*$) and closures (threshold $A_R^*$). $W$ is proportional to the social value of the bank, net of auditing costs

---

14
In Figure 3 we represent the optimal combination of inspection and closure thresholds by \((A^*_R, A^*_I)\). It is obtained by maximizing the social value of the bank:

\[
V(A, A_R) = A - \gamma + [\gamma - (1 - \lambda)A_R] \left( \frac{A}{A_R} \right)^{-a},
\]

net of the expected present value of auditing costs:

\[
C(A, A_R, A_I) = E \left( \int^{\tau_R} \xi \mathbb{I}_{A_t \leq A_I} e^{-rt} dt \right) dt,
\]

where the auditing cost \(\xi\) is only incurred when the asset value of the bank lies in the interval \([A_R, A_I]\), and \(\tau_R\) denotes the first time where \(A_t = A_R\) (closure time). It can be shown that the expected present value of auditing costs \(C\) is proportional to \(A^{-a}\):

\[
C(A, A_R, A_I) = A^{-a} \varphi(A_R, A_I).
\]

**Proposition 5**: The optimal value of the closure threshold \(A^*_R\) and the inspection threshold \(A^*_I\) can be obtained by maximizing

\[
W(A_R, A_I) \equiv \left( \gamma - (1 - \lambda)A_R \right) A_R^a - \varphi(A_R, A_I),
\]

under the incentive compatibility constraint (15).

We have that \(A^*_R < A^*_R^0\), which means that Prompt Corrective Action (PCA) allows to reduce capital requirements. When auditing costs are small, \(A^*_R\) becomes close to the first best closure threshold \(A_{FB}\), which means that PCA also reduces the intensity of the time consistency problem of bank supervisors.

Proposition 5 provides an illustration of the substitutability between Pillar 2 (supervisory action) and Pillar 1 (capital requirements). This was already a feature of Merton’s (1978) model, where frequency of bank examinations could be substituted to more stringent capital requirements. However here, the introduction of Pillar 3 (market discipline) changes the picture: the intensity of regulation can be modulated according to market information (in the spirit of the Prompt Corrective Action provisions of FDICIA) and symmetrically, supervisors can be forced to intervene when market signals reveal the distress of a bank (so that forbearance becomes more costly to supervisors or politicians).

### 7 Conclusion

This paper develops a formal model of banking regulation that permits to analyze the interactions between the 3 pillars of Basel 2. It differs from previous models in several important ways:
- it is a dynamic model, where solvency regulations are interpreted as closure thresholds, rather than complex tools intended to correct the mispricing of deposit insurance;

- the justification of regulation is not (primarily) to prevent asset substitution by banks (deposits are not subsidized in our model) but rather to prevent shirking by the managers of undercapitalized banks;

- bank supervisors can use market information as a useful complement to the information provided by bank examinations. Thus they can save on scarce supervisory resources and allocate them in priority to the banks in distress;

- the returns on banks’ assets are endogenous, since they depend on the monitoring decisions of bankers.

Although very simple, this model allows a formal analysis of the interactions between the 3 pillars of Basel 2. In particular we show in Proposition 3 that mandatory subdebt (direct market discipline) may, under some restrictions, allow regulators to decrease capital requirements. More importantly, we show that market discipline and supervisory action are complements rather than substitutes (Proposition 4 and 5): one cannot work well without the other.

In terms of policy implications, our theory points toward a serious rebalancing of the 3 pillars of Basel 2. The initial motivation of Basel 1 was to make a level playing field for international banking, given that the large banks of some countries could take enormous risks without having much capital, benefiting from implicit guarantees by their governments. So the fundamental idea behind the Cooke ratio was harmonization, i.e. to set a uniform standard for internationally active banks. It turns out that this Cooke ratio (or more generally the risk-based capital methodology), although imperfect, revealed extremely useful as an instrument for measuring bank risk. This is probably why it was rapidly applied (with minor changes) by the regulatory authorities of many countries within their jurisdictions (although it was initially designed for large, internationally active, banks).

Probably traumatized by the harsh critiques addressed to the crudeness of the Cooke Ratio, the Basel Committee began a process of complexification, alternating new proposals and consultation periods with the banking industry (Basel Committee, 1999, 2001 and 2003). The outcome of this long process is an extremely complex instrument (the McDonough ratio), that results from intense bargainings with large banks, and will probably be never implemented as such. It is as if we were back to the old days of banking supervision, where bankers used to comply with the instructions of paternalistic supervisors, in exchange for protection from competition by new entrants.
I claim that banking authorities should instead keep arm’s length relationships with bankers and that scarce supervisory resources should be used in priority to control strictly the behavior of banks in distress, rather than trying to implement an extremely complex regulation that will ultimately be bypassed in some way or another by the largest or most sophisticated banks. By contrast, there is an urgent need (once again) to make a level playing field in international banking. The development of Large and Complex Banking Organizations with multinational activities implies that supervisory authorities of different countries need urgently to harmonize their institutional practices. Market discipline can provide a very useful tool for defining an harmonized and clear mandate for banking authorities all across the world, in an attempt to eliminate political pressure and regulatory forbearance. This should be priority number one of the Basel Committee.
Mathematical Appendix

We derive in this section the mathematical formulas used in the text.

1) First best value of the bank

The value of the bank (when assets are monitored) equals the expected present value of future cash flows $\beta A_t$, net of the monitoring cost $r \gamma$, until the stopping time $\tau_L$ (first time $t$ where $A_t$ hits the liquidation threshold $A_L$), where the bank is liquidated. The formula is:

$$ V = E \left[ \int_0^{\tau_L} e^{-rt} (\beta A_t - r \gamma) dt + \frac{\lambda A_L}{r - \mu} \right]. \quad (A.1) $$

Using classical formulas (see for example Dixit, 1993, or Karlin and Taylor, 1981) we obtain:

$$ V = A_t - \gamma + \{ \gamma - (1 - \lambda)A_L \} \left( \frac{A}{A_L} \right)^{-a}, \quad (A.2) $$

where $A$ is the current value of $A_t$, and $a$ is the positive root of the quadratic equation:

$$ \frac{1}{2} \sigma^2 x^2 - \mu x = r. \quad (A.3) $$

2) Value of the bank’s equity

In the absence of regulation, equityholders choose the liquidation threshold that maximizes the value of their equity. Using the same classical formulas as for establishing (A.2) we obtain:

$$ E(A) = (A - \gamma - D) + (D + \gamma - A_L) \left( \frac{A}{A_L} \right)^{-a}. \quad (A.4) $$

As for the total value of the bank (Formula A.2) the second term is an option value that is maximized when

$$ A_L = A_E \equiv \frac{a}{a + 1}(D + \gamma). \quad (A.5) $$

At this threshold, the value of the bank’s equity has an horizontal tangent (as represented in Figure 2 in the text):

$$ E'(A_E) = 0. \quad (A.6) $$
If equity holders decide to stop monitoring, the dynamics of assets value becomes

$$\frac{dA}{A} = (\mu - \Delta \mu) dt + \sigma dW,$$

but they save the monitoring cost $r\gamma$. Shirking becomes optimal for equity holders whenever the instantaneous loss of equity value $E'(A) A \Delta \mu$ is less than this monitoring cost. Because $E'(A_E) = 0$ (see condition (A.5)), this condition is always satisfied in the neighborhood of the liquidation point. However we have to check that this incentive constraint binds after the bank becomes insolvent. This is true whenever $\lambda A_E \leq D$, or

$$\lambda a(D + \gamma) \leq (a + 1) D,$$

which is equivalent to the condition of Proposition 1, namely

$$\frac{\gamma}{D} \leq \frac{a + 1}{\lambda a} - 1.$$

This ends the Proof of Proposition 1.

3) Minimum capital ratio

Suppose bank regulators impose a closure threshold $A_R \leq \frac{D}{\gamma}$: if the banks’ asset value hits $A_R$, the bank is liquidated and shareholders receive nothing. By an immediate adaptation of formula (A.4), shareholders value becomes:

$$E(A) = A - \gamma - D + (D + \gamma - A_R) \left( \frac{A}{A_R} \right)^{-a}. \tag{A.7}$$

The condition for eliminating shirking is:

$$\forall A \geq A_R, \quad E'(A) A \Delta \mu \geq \gamma r. \tag{A.8}$$

Using (A.7), this is equivalent to:

$$\forall A \geq A_R, \quad A - a(D + \gamma - A_R) \left( \frac{A}{A_R} \right)^{-a} \geq \frac{\gamma r}{\Delta \mu}. \tag{A.9}$$

Provided $A_R \leq \gamma + D$ (this will be checked ex post), the left-hand side of (A.9) is increasing in $A$, therefore condition (A.9) is equivalent to:

$$A_R(1 + a) - a(D + \gamma) \geq \frac{\gamma r}{\Delta \mu},$$

or:

$$A_R \geq A_R^O \equiv \frac{a(D + \gamma) + \frac{\gamma r}{\Delta \mu}}{a + 1}. \tag{A.10}$$

19
$A_R^O$ represents the minimum asset value that preserves the incentives of the banker. The associated capital ratio is:

$$
\rho^R = \frac{A_R^O - D}{A_R^O} = \frac{\gamma(a + \frac{r}{\sqrt{\mu}}) - D}{a(D + \gamma) + \frac{r}{\sqrt{\mu}}},
$$

4) Subordinated debt

Consider now that the bank issues a volume $B$ of subordinated bonds, paying a coupon $cB$ per unit of time, and randomly renewed with frequency $m$. The market value of these bonds $B(A)$ (as a function of the bank’s asset value) satisfies the differential equation

$$
rb(A) = cB + m(B - B(A)) + \mu AB'(A) + \frac{1}{2} \sigma^2 A^2 B''(A),
$$

(A.11)

with the boundary conditions:

$$
B(A_L) = 0 \quad \text{and} \quad B(+\infty) = \frac{cB}{r}.
$$

The solution of this equation is:

$$
B(A) = \frac{cB}{r} \left[ 1 - \left( \frac{A}{A_L} \right)^{-a(m)} \right],
$$

(A.12)

where $a(m)$ is the positive root of the quadratic equation:

$$
\frac{1}{2} \sigma^2 x^2 - \mu x = r + m.
$$

(A.13)

Comparing with (A.3), we see immediately that $a(0) = a$. Moreover, (A.13) shows that $a(m)$ increases with $m$.

The value of equity becomes:

$$
E(A, B) = A - \gamma - D - \frac{c}{r} B + (D + \gamma - A_L) \left( \frac{A}{A_L} \right)^{-a} + \frac{c}{r} B \left( \frac{A}{A_L} \right)^{-a(m)}.
$$

(A.14)

5) Cost of auditing

By definition the expected present value of auditing costs is defined by

$$
C(A, A_R, A_I) = E \left[ \int_0^{\tau_R} \xi I_{A_t \leq A_I} e^{-rt} dt | A \right],
$$

where $\tau_R$ is the first time where $A_t$ hits the closure threshold $A_R$. By usual arguments (see Dixit 1993) one can establish that $C$ satisfies the following differential equation

$$
rC = \mu AC''(A) + \frac{1}{2} \sigma^2 A^2 C''(A) \quad A \geq A_I,
$$

20
with the limit condition

\[ C(+\infty) = 0. \]

Therefore \( C(A) = kA^{-a} \), where \( a \) is (as before) the positive solution of the equation:

\[ r = -\mu x + \frac{1}{2} \sigma^2 x^2, \]

and \( k \) is a constant that depends on \( A_R \) and \( A_I \):

\[ k = \varphi(A_R, A_I). \]
References


Merton, R., (1977), ”An analytic derivation of the cost of deposit insurance and loan guarantees: an application of modern option pricing theory”, *Journal of Banking and Finance*, 1, 3-11.

Merton, R., (1978), ”On the cost of deposit insurance when there are surveillance costs”, *Journal of Business*, 51, 439-452.


