The Three Pillars of Basel II: Optimizing the Mix

by

Jean-Paul DECAMPS\textsuperscript{1}
Jean-Charles ROCHET\textsuperscript{2}
Benoît ROGER\textsuperscript{3}

Revised Version
April 2003

Abstract

The on-going reform of the Basel Accord relies on three “pillars”: a new capital adequacy requirement, supervisory review and market discipline. This article develops a simple continuous-time model of commercial banks’ behavior where the articulation between these three instruments can be analyzed. We study the conditions under which market discipline can reduce the minimum capital requirements needed to prevent moral hazard. We also discuss regulatory forbearance issues.

Acknowledgment: We benefited from the comments of Sudipto Bhattacharya, Bernard Dumas, Charles Goodhart, Henri Pagès, Rafael Repullo, Joao Santos, Elu Von Thadden, Stéphane Villeneuve, and an anonymous referee, as well as seminar participants at the Chicago Fed, the Financial Markets Group (L.S.E.), the B.I.S., and I.N.S.E.A.D. The usual disclaimer applies.

\textsuperscript{1}Université des Sciences Sociales(GREMAQ), Toulouse, France.
\textsuperscript{2}London School of Economics and Université des Sciences Sociales(IDEI), Toulouse, France. Corresponding author. Email: rochet@cict.fr.
\textsuperscript{3}Université des Sciences Sociales(GREMAQ), Toulouse, France.
1 Introduction

The on-going reform of the Basel Accord ¹ relies on three “pillars”: capital adequacy requirements, supervisory review and market discipline. Yet, the articulation between these three instruments is far from being clear. On the one hand, the recourse to market discipline is rightly justified by common sense arguments about the increasing complexity of banking activities, and the impossibility for banking supervisors to monitor in detail these activities. It is therefore legitimate to encourage monitoring of banks by professional investors and financial analysts as a complement to banking supervision. Similarly, a notion of gradualism in regulatory intervention is introduced (in the spirit of the reform of US banking regulation, following the FDIC Improvement Act of 1991).² It is suggested that commercial banks should, under “normal circumstances”, maintain economic capital way above the regulatory minimum and that supervisors could intervene if this is not the case. Yet, and somewhat contradictorily, while the proposed reform states very precisely the complex refinements of the risk-weights to be used in the computation of this regulatory minimum, it remains silent on the other intervention thresholds.

It is true that the initial accord (Basel, 1988) has been severely criticized for being too crude,³ and introducing a wedge between the market assessment of asset risks and its regulatory counterpart.⁴ However, it seems contradictory to insist so much on the need to “enable early supervisory intervention if capital does not provide a sufficient buffer against risk” and to remain silent on the threshold and form of intervention, while putting so much effort on the design of risk weights. Similarly, nothing very precise is said (apart from the need for”increased transparency” !)about the way to implement Pillar 3 (market discipline) in practice.⁵

A possible explanation for this inbalance between Pillar 1 and the other two pillars is that most of the formal analyses of banks capital regulation rely on static models,

¹The Basel Accord, elaborated in July 1988 by the Basel Committee on Banking Supervision (BCBS) required internationally active banks from the G10 countries to hold a minimum total capital equal to 8% of risk-adjusted assets. It was later amended to cover market risks. It is currently being revised by the BCBS, who has released for comment a proposal of amendment, commonly referred to as Basel II (Basel Committee, 1999, 2001).

²The FDIC Improvement Act of 1991 requires that each US bank be placed in one of five categories based on its regulatory capital position and other criteria (CAMELS ratings). Undercapitalized banks are subject to increasing regulatory intervention as their capital ratios deteriorate. This prompt corrective action (PCA) doctrine is designed to limit supervisory forbearance. Jones and King (1995) provide a critical assessment of PCA. They suggest that the risk weights used in the computation of capital requirements are inadequate.

³Jones (2000) also criticizes the Basel Accord by showing how banks can use financial innovation to increase their reported capital ratios without truly enhancing their soundness.

⁴See our discussion of the literature in Section 2.

⁵In particular, in spite of the existence of very precise proposals by US economists (Evanoff and Wall (2000), Calomiris (1998), see also the discussion in Bliss (2001)) for mandatory subordinated debt, these proposals are not discussed in the Basel 2 project.
where capital requirements are used to curb banks’ incentives for excessive risk-taking and where the choice of risk weights is fundamental. However, as suggested by Hellwig (1998), a static framework fails to capture important intertemporal effects. For example, in a static model, a capital requirement can only have an impact on banks’ behaviour if it is binding. In practice however, capital requirements are binding for a very small minority of banks and yet seem to influence the behavior of other banks. Moreover, as suggested by Blum (1999), the impact of more stringent capital requirements may sometimes be counter intuitive, once intertemporal effects are taken into account. The modeling cost is obviously additional complexity, due in particular to transitory effects. In order to minimize this complexity, we will assume here a stationary liability structure, and rule out those transitory effects. Also for simplicity, we will only consider one type of assets, allowing to derive a Markov model of banks’ behavior with only one state variable: the cashflows generated by the bank’s assets (or, up to a monotonic transformation, the bank’s capital ratio).

In this paper, we adopt the view, consistent with the approach of Dewatripont and Tirole (1994) that capital requirements should be viewed as intervention thresholds for banking supervisors (acting as representatives of depositors’ interest) rather than complex schemes designed to curb banks’ asset allocation. This means that we will not discuss the issue of how to compute risk weights (it has already received a lot of attention in the recent literature), but focus instead on what to do when banks do not comply with capital requirements, a topic that seems to have been largely neglected.

We build on a series of recent articles that have adapted continuous time models used in the corporate finance literature to analyze the impact of the liability structure of firms on their choices of investment and on their overall performance. We extend this literature by incorporating features that we believe essential to capture the specificities of commercial banks.

We model banks as “delegated monitors” à la Diamond (1984) by considering that banks have the unique ability to select and monitor investments with a positive net present value and finance them in large part by deposits. Liquidation of banks is costly because of the imperfect transferability of banks’ assets. Also, profitability of these investments requires costly monitoring by the bank. Absent the incentives for the banker to monitor, the net present value of his investments becomes negative. We show that these incentives are absent precisely when the bank is insufficiently capitalized. Thus incentive compatibility conditions create the need for the regulator, acting on behalf of depositors, to limit banks’ leverage and to impose closure well before the net present value of the bank’s assets becomes negative. This is the justification for capital requirements in our model.
Notice that there are two reasons why the Modigliani Miller theorem is not valid in our model: the value of the bank is indeed affected both by closure decisions and by moral hazard on investment monitoring by bankers. Closure rules, i.e. capital requirements, optimally trade-off between these two imperfections. However, these capital requirements give rise to a commitment problem for supervisors: from a social welfare perspective, it is almost always optimal to let a commercial bank continue to operate, even if this bank is severely undercapitalized. Of course, this time inconsistency problem generates bad incentives for the owners of the bank from an ex-ante point of view, unless the bank’ supervisors find a commitment device, preventing renegotiation.

The rest of the paper is organized as follows. After a brief review of the literature in Section 2, we describe our model in Section 3. In Section 4 we provide the justification for solvency regulations: a minimum capital requirement is needed to prevent insufficiently capitalized banks from shirking . In Section 5 we introduce market discipline through compulsory subordinated debt. We show that under certain circumstances it may reduce the minimum capital requirement. Section 6 analyses supervisory action. We show that direct market discipline is only effective when the threat of bank closures by supervisors is credible. In this case, indirect market discipline can also be useful in allowing supervisors to implement gradual interventions.

2 Related Literature

We will not discuss in detail the enormous literature on the Basel Accord and its relation with the “credit crunch” (good discussions can be found in Thakor (1996), Jackson et al. (1999), Santos (2000)). Let us briefly mention that most of the theoretical literature (e.g., Furlong and Keeley (1990), Kim and Santomero (1988), Koehn and Santomero (1980), Rochet (1992), Thakor (1996)) has focused on the distortion of banks’ assets allocation that could be generated by the wedge between market assessment of asset risks and its regulatory counterpart in Basel I. The empirical literature (e.g. Bernanke and Lown (1991); see also Thakor (1996), Jackson et al. (1999) and the references therein) has tried to relate these theoretical arguments to the spectacular (yet apparently transitory) substitution of commercial and industrial loans by investment in government securities in US banks in the early 1990s, shortly after the implementation of the Basel Accord and FDICIA.\footnote{Peek and Rosengren (1995) find that the increase in supervisory monitoring had also a significant impact on bank lending decisions, even after controlling for bank capital ratios.} \footnote{Blum and Hellwig (1995) analyze the macroeconomic implications of bank capital regulation.}

Even if these authors seem to have established a positive correlation between bank capital and commercial lending, causality can only be examined in a dynamic framework. Blum (1999) is one of the first theoretical papers to analyze the consequences of more
stringent capital requirements in a dynamic framework. He shows that more stringent capital requirements may paradoxically induce an increase in risk taking by the banks who anticipate having difficulty meeting these capital requirements in the future.

Hancock et al. (1995) study the dynamic response to shocks in the capital of US banks using a Vector Auto Regressive framework. They show that US banks seem to adjust their capital ratios must faster than they adjust their loans portfolios. Furfine (2001) extends this line of research by building a structural dynamic model of banks behavior, which is calibrated on data from a panel of large US banks on the period 1990-97. He suggests that the credit crunch cannot be explained by demand effects but rather by the raise in capital requirements and/or the increase in regulatory monitoring. He also uses his calibrated model to simulate the effects of Basel II and suggests that its implementation would not provoke a second credit crunch, given that average risk weights on good quality commercial loans will decrease if Basel II is implemented.

Our objective here is to design a tractable dynamic model of bank behavior where the articulation between the three pillars of Basel II can be analyzed. Our model builds on two strands of the literature:

- Corporate finance models à la Leland and Toft (1996) and Ericsson (2000), who analyze the impact of debt maturity on asset substitution and firm value;
- Banking models à la Merton (1977), Fries et al. (1997), Bhattacharya et al. (2000), Milne and Whalley (2001) who analyze the impact of solvency regulations and supervision intensity on the behavior of commercial banks.

Let us briefly summarize the main findings of these articles.

Leland and Toft (1996) investigate the optimal capital structure which balances the tax benefits coming with debt and bankruptcy costs. They extend Leland (1994) by considering a coupon bond with finite maturity $T$. They maintain the convenient assumption of a stationary debt structure by assuming a constant renewal of this debt at rate $m = \frac{1}{T}$. Leland and Toft (1996) are able to obtain closed form (but complex) formulas for the value of debt and equity. In addition, using numerical simulations, they show that risk shifting disappears when $T \to 0$, in conformity with the intuition that short term debt allows to discipline managers.8

Ericsson (2000) and Leland (1998) also touch on optimal capital structure, but are mainly concerned with the asset substitution problem arising when the managers of a firm can modify the volatility of its assets’ value. They show how the liability structure

influences the choice of assets’ volatility by the firm’s managers. Both consider a perpetual debt but Ericsson (2000) introduces a constant renewal rate which serves as a disciplining instrument.

Mella-Barral and Perraudin (1997) characterise the consequences of the capital structure on an abandonment decision. They obtain an underinvestment (i.e. premature abandonment) result. This comes from the fact that equityholders have to inject new cash in the firm to keep it as an ongoing concern. Similarly, Mauer and Ott (1998) consider the investment policy of a leveraged company and also exhibit an underinvestment result for exactly the same reason. These papers thus offer a continuous time version of the debt overhang problem first examined in Myers (1977): the injection of new cash by equityholders creates a positive externality on debtholders’ claims and the continuation (or expansion) decisions are under-optimal because equityholders do not internalize this effect. Anderson and Sundaresan (1996) and Mella-Barral (1999) elaborate on this aspect by studying the impact of possible renegotiation between equityholders and debtholders. They also allow for the possibility of strategic default.

In the other strand of the literature, Merton (1977, 1978) is the first to use a diffusion model for studying the behavior of commercial banks. He computes the fair pricing of deposit insurance in a context where supervisors can perform costly audits. Fries et al. (1997) extend Merton’s framework, by introducing a withdrawal risk on deposits. They study the impact of the regulatory policy of bank closures on the fair pricing of deposit insurance. The optimal closure rule has to trade-off between monitoring costs and costs of bankruptcy. Under certain circumstances, the regulator may want to let the bank continue even when equity-holders have decided to close it (underinvestment result).

Following Leland (1994), Bhattacharya et al. (2002) derive closure rules that can be contingent on the level of risk chosen by the bank. Then they examine the complementarity between two policy instruments of bank regulators: the level of capital requirements and the intensity of supervision. In the same spirit, Dangl and Lehar (2001) mix random audits as in Bhattacharya et al. (2002) with risk shifting possibilities as in Leland (1998) so as to compare the efficiency of Basel Accords (1988) and VaR regulation. They show that VaR regulation is better, since it reduces the frequency of audits needed to prevent risk shifting by banks.

Calem and Rob (1996) design a dynamic (discrete time) model of portfolio choice, and analyse the impact of capital based-premia when regulatory audits are perfect. They show that regulation may be counterproductive: a tightening in capital requirement may lead to an increase in the risk of the portfolios chosen by banks, and similarly, capital-based premia may sometimes induce excessive risk taking by banks. However, this never happens when capital requirements are stringent enough.
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. Milne and Whalley (2001) develop a model where banks can issue subsidized deposits without limit in order to finance their liquidity needs. The social cost of these subsidies is limited by the threat of regulatory closure. Milne and Whalley study the articulation between two regulatory instruments: the intensity of costly auditing and the level of capital requirements. They also allow for the possibility of banks recapitalization. They show that banks’optimal strategy is to hold an additional amount of capital (above the regulatory minimum) used as a buffer against future solvency shocks. This buffer reduces the impact of solvency requirements.

Finally, Pagès and Santos (2001) analyze optimal banking regulations and supervisory policies according to whether or not banking authorities are also in charge of the deposit insurance fund. If this is the case, Pagès and Santos show that supervisory authorities should inflict higher penalties on the banks who do not comply with solvency regulations, but should also reduce the frequency of regulatory audits.

We now move on to the description of our model.

3 The Model

Following Merton (1974), Black and Cox (1976) and Leland (1994), we model the cash-flows $x$ generated by the bank’s assets by a diffusion process:

$$
\frac{dx}{x} = \mu_G dt + \sigma_G dW.
$$ (1)

We also assume all agents are risk neutral with a discount rate $r > \mu_G$.

Equation (1) is only satisfied if the bank monitors its assets. Monitoring has a fixed (non monetary) cost per unit of time, equivalent to a continuous monetary outflow $rb$. This $b$ can thus be interpreted as the present value of the cost of monitoring the bank’s assets forever. In the absence of monitoring, the cash-flows dynamics satisfies instead:

$$
\frac{dx}{x} = \mu_B dt + \sigma_B dW,
$$ (2)

If monitoring cost has also a variable component, it can be substracted from $\mu_G$. This monitoring cost captures the efforts that bankers have to exert in order to extract adequate repayments from borrowers, or alternatively the foregone private benefits that could have been obtained by related lending. Being non monetary, this cost does not appear in accounting values but it does affect the (market) value of equity for bankers.

For simplicity, we assume that the bad technology choice is irreversible: once the bank has started "shirking", the dynamics of $x$ is forever given by equation (2). Reversible choices would lead to similar results, with slightly more complicated formulas. Reversibility would also complicate our analysis of regulatory forbearance in Section 6.
where \( B \) stands for the “bad” technology (and \( G \) for the “good” technology) and \( \mu_B = \mu_G - \Delta \mu \leq \mu_G \) and \( \sigma^2_B = \sigma^2_G + \Delta \sigma^2 \geq \sigma^2_G \). For technical reasons, we also assume that \( \sigma^2_G < \frac{1}{2}(\mu_G + \mu_B) \).

Notice that when \( \Delta \sigma^2 = 0 \), we have the classical first order stochastic dominance (pure effort) problem. When \( \Delta \sigma^2 > 0 \), there is also a risk shifting component.

If the bank is closed, the bank’s assets are liquidated for a value \( \lambda x \) (i.e. that is proportional to the current value of cash flows \(^{11}\)). \( \lambda \) is given exogenously, and satisfies:

\[
\frac{1}{r - \mu_B} < \lambda < \frac{1}{r - \mu_G}. \quad (3)
\]

The first inequality means that closure is always preferable to the "bad" technology:

\[
E_{x_0} \left[ \int_0^{+\infty} e^{-rt} x_t dt \right| \text{bad technology} ] = \frac{x_0}{r - \mu_B} < \lambda x_0.
\]

The second inequality captures the assumption that outsiders are only able to capture some fraction \( \lambda (r - \mu_G) < 1 \) of the future cash flows delivered by the bank’s assets. However, due to the fixed monitoring cost \( r b \), liquidation is optimal when \( x_0 \) is small. Indeed, the net present value of a bank who continuously monitors its assets is:

\[
E_{x_0} \left[ \int_0^{+\infty} e^{-rt} x_t - rb dt \right| \text{good technology} ] = \frac{x_0}{r - \mu_G} - b,
\]

so the “good” technology dominates closure whenever \( x_0 \) is not too small:

\[
\frac{x_0}{r - \mu_G} - b > \lambda x_0 \quad \Leftrightarrow \quad x_0 > \frac{b}{\nu_G - \lambda},
\]

where \( \nu_G = \frac{1}{r - \mu_G} > \lambda \),

while we denote by analogy:

\( \nu_B = \frac{1}{r - \mu_B} < \lambda \).

In the absence of a closure threshold (i.e. assuming that banks continue forever), the surpluses generated by the good (\( G \)) and the bad (\( B \)) technologies would be as represented in Figure 1 below.

\(^{11}\)Mella-Barral and Perraudin (1997) assume instead a constant liquidation value.

\(^{12}\)Genotte and Pyle (1991) were the first to analyze capital regulations in a framework where banks have an explicit monitoring role and make positive NPV loans. In some sense, our paper can be viewed as a dynamic version of Genotte and Pyle (1991).

\(^{13}\)This implies that banks’ assets are not traded and thus markets are not complete. In a complete markets framework, the moral hazard problem can be solved by risk-based deposit insurance premia and capital regulation becomes redundant.
The economic surplus generated by the good technology is therefore positive when $x$ is larger than the NPV threshold $\frac{b}{V_G - \lambda}$, while the surplus generated by the bad technology is always negative. We now introduce a closure decision, determined by a liquidation threshold $x_L$.

Assuming for the moment that the bank always monitors its assets ("good technology"), the value of these assets $V_G(x)$ is thus determined by the liquidation threshold $x_L$, below which the bank is closed:

$$V_G(x) = E_x \left[ \int_0^{\tau_L} e^{-rt}(x_t - rb)dt + e^{-r\tau_L} \lambda x_L \right],$$

(4)

where $\tau_L$ is a random variable (stopping time), defined as the first instant where $x_t$ (defined by (1)) equals $x_L$, given $x_0 = x$. 

Figure 1: Economic surpluses generated by the good (G) and the bad (B) technologies.
Using standard formulas, we obtain:

\[ V_G(x) = \nu_G x - b + \{ b - (\nu_G - \lambda) x_L \} \left( \frac{x}{x_L} \right)^{1 - a_G}, \quad (5) \]

where

\[ a_G = \frac{1}{2} + \frac{\mu_G}{\sigma_G^2} + \sqrt{\left( \frac{\mu_G}{\sigma_G^2} - \frac{1}{2} \right)^2 + \frac{2r}{\sigma_G^2}} > 1. \quad (6) \]

The continuation value of the bank is thus equal to the net present value of perpetual continuation \((\nu_G x - b)\) plus the option value associated to the irreversible closure decision at threshold \(x_L\). Interestingly this option value is proportional to \(x^{1-a_G}\), thus it is maximum for a value of \(x_L\) that does not depend on \(x\), namely

\[ x_{FB} = \frac{b}{(\nu_G - \lambda)} \frac{a_G - 1}{a_G}. \quad (7) \]

**Proposition 1**: The first best closure threshold of the bank is the value of \(x_L\) that maximizes the option value associated to the irreversible closure decision. This value is equal to \(x_{FB} = \frac{b}{(\nu_G - \lambda)} \frac{a_G - 1}{a_G}\). It is smaller than the NPV threshold \(\frac{b}{\nu_G - \lambda}\).

The continuation value of the bank as a function of \(x\) (i.e. \(V_G(x) - \lambda x\)) is represented below for different values of \(x_L\):

\[ \text{See for instance Karlin and Taylor (1981).} \]
Figures 2: The continuation value of the bank for different closure thresholds:

- \( x^A_L \) corresponds to excessive continuation \((V'_{G}(x^A_L) < \lambda)\),
- \( x^B_L \) corresponds to premature closure \((V'_{G}(x^B_L) > \lambda)\),
- \( x_{FB} \) corresponds to the optimal threshold \((V'_{G}(x_{FB}) = \lambda)\),
- \( \frac{b}{(\nu_G - \lambda)} \) corresponds to the positive NPV threshold.

We now introduce the second characteristic feature of commercial banking, namely deposit finance: a large fraction of the bank’s liabilities consists of insured deposits,\(^{15}\) with a volume normalized to one. For the moment, we assume that these deposits are the only source of outside funds for the bank (we later introduce subordinated debt) and that issuing equity is prohibitively costly.\(^{16}\) In the absence of public intervention,\(^{17}\) liquidation of the bank occurs when the cash flows \( x \) received from its assets are insufficient to repay the interest \( r \) on deposits. In this case, the liquidation threshold is thus:

\[
x_L = r.
\]  

\(^{15}\)For simplicity, we assume that these are long term deposits. It would be easy to introduce a constant frequency of withdrawals, as in our treatment of subordinated debt in Section 5.

\(^{16}\)Bhattacharya et al. (2002) make instead the assumption that the bank can costlessly issue new equity. In that case, the closure threshold is chosen by stockholders so as to maximize equity value. Milne and Whalley (2001) make the intermediate assumption that new equity issues entail an exogenous fixed cost.

\(^{17}\)Public intervention can consist either of liquidity assistance by the Central Bank, or on the contrary closure by the banking supervision authorities. This is analyzed in the next sections.
We also assume that when this liquidation takes place, the book value of the bank equity (which, in our model, is equal to the book value of assets $\nu_G x$ minus the nominal value of deposits) is still positive:

$$\nu_G r > 1,$$

but liquidation does not allow to repay all deposits:

$$\lambda r < 1.$$ 

Condition (9) captures the fact that, in the absence of liquidity assistance by the central bank (introduced in Section 6) solvent banks may be illiquid.\(^{18}\) Condition (10) ensures that deposits are risky.

The $PV$ of deposits is computed easily:

$$D_G(x) = 1 - (1 - \lambda x_L) \left( \frac{x}{x_L} \right)^{1-a_G},$$

leading to the market value of equity:

$$E_G(x) = V_G(x) - D_G(x),$$

or

$$E_G(x) = \nu_G x - 1 + (b + 1 - \nu_G x_L) \left( \frac{x}{x_L} \right)^{1-a_G}.$$ 

Notice that since $\lambda x_L < 1$, deposits are risky. As a result, the $PV$ of deposits $D(x)$ is less than their nominal value 1, the difference corresponding to the liability of the Deposit Insurance Fund.\(^{19}\) If instead the bank ceases to monitor its assets, the value of equity becomes, by a simple adaptation of the above formula (replacing $\nu_G$ by $\nu_B$ and $b$ by zero):

$$E_B(x) = \nu_B x - 1 + (b + 1 - \nu_B x_L) \left( \frac{x}{x_L} \right)^{1-a_B},$$

where

$$a_B = \frac{1}{2} + \frac{\mu_B}{\sigma_B^2} + \sqrt{\left( \frac{\mu_B}{\sigma_B^2} - \frac{1}{2} \right)^2 + \frac{2 \rho}{\sigma_B^2}}.$$ 

By comparing the value of equity in the two formulas, it is easy to see that in general $E_B(x) > E_G(x)$ for $x$ in some interval $[x_L, x_S]$, as suggested by the following figure:

\(^{18}\)This assumption is in line with Bagehot’s doctrine for a Lender of Last Resort (see for example Rochet and Vives (2002) for a recent account of this doctrine). In our model, it guarantees that optimal capital requirements are positive. However it is not crucial: even if it is not satisfied, optimal capital requirements are positive if $b$ is large enough (see below).

\(^{19}\)This liability is covered by an insurance premium $1 - D(x_0)$ paid initially by the bank. We could also introduce a flow premium, paid in continuous time, as in Fries et al. (1997).
Figure 3: Comparing equity values under good and bad technology choices.

4 The Justification of Solvency Requirements

Figure 3 illustrates the basic reason for imposing a capital requirement in our model: as long as $E_G'(x_L) < E_B'(x_L)$, there is a region $[x_L, x_S]$ where, in the absence of outside intervention, the bank “shirks” (i.e. chooses the bad technology) which reduces social welfare, and ultimately provokes failure, the cost being borne by the Deposit Insurance Fund (DIF). As shown by Proposition 2 below, this happens when the monitoring cost $b$ is not too small. In order to avoid shirking, banking authorities (which could be the Central Bank, a Financial Services Authority or the DIF itself) set a regulatory closure threshold $x_R$ below which the bank is closed. In practice, this closure threshold can be implemented by a minimal capital requirement. Indeed, the book value of equity is equal to the book value of assets $\nu_G x$, minus the nominal value of deposits, which we have normalized to 1. The solvency ratio of the bank is thus:

$$\rho = \frac{\nu_G x - 1}{\nu_G x}.$$

This is an increasing function of $x$. A continuation rule $x \geq x_R$ is thus equivalent to a minimum capital ratio

$$\rho \geq \frac{\nu_G x_R - 1}{\nu_G x_R} \overset{\text{def}}{=} \rho_R.$$ 

Proposition 2: When $b$ is not too small, a solvency regulation is needed to prevent insufficiently capitalized banks from shirking. The second best closure threshold (associated
with the optimal capital ratio) is the smallest value \( x_R \) of the liquidation threshold such that shirking disappears. It is given by:

\[
x_R = \frac{(a_G - 1)b + a_G - a_B}{a_G \nu_G - a_B \nu_B}.
\] (15)

Regulation is needed whenever \( x_R > x_L \) which is equivalent to

\[
b > \hat{b} = \frac{r[a_G \nu_G - a_B \nu_B] - (a_G - a_B)}{a_G - 1}.
\] (16)

**Proof:** See Appendix.

Notice that, when regulation is needed \( (x_R > x_L = r) \), the implied solvency ratio \( \rho_R \) is always positive:

\[
\rho_R = \frac{\nu_G x_R - 1}{\nu_G x_R} > 0.
\]

This is because we have assumed \( \nu_G r > 1 \), which means that, in the absence of public intervention, banks become illiquid before they become insolvent.\(^{20}\)

Notice also that when \( b \) is very large, the liquidation value of the bank \( \lambda x_R \) becomes greater than the nominal value of deposits (normalized to 1) and deposits become riskless. In this case the incentives of banks’ stockholders are not distorted by the limited liability option: they optimally decide to close the bank when \( x \) hits the first best threshold \( x_{FB} \) and the moral hazard constraint does not bind. We focus on the more interesting set of parameters values for which undercapitalized banks have indeed incentives to shirk.

### 5 Market Discipline

There are several reasons why market discipline can be useful. First it can produce additional information that the regulator can exploit (this is usually referred to as “indirect” market discipline). Consider for example a set up à la Merton (1978) or Bhattacharya et al. (2002) where \( x_t \) is only observed through costly and imperfect auditing. As a result, there is a positive probability that the bank may continue to operate in the region \([x_L, x_R]\) (because undetected by banking supervisors). If shirking is to be deterred, a more stringent capital requirement (i.e. a higher \( x_R \)) has to be imposed, to account for imperfect auditing (see Bhattacharya et al. (2002) for details). In such a context, imposing the bank to issue a security (say subordinated debt) whose pay-off is conditional on \( x_t \), and

\(^{20}\) Even without this assumption, \( \rho_R \) is positive whenever \( b \) is large enough. This is in line with most corporate finance models with moral hazard (for example, Holmström and Tirole (1997)): when the cost of effort (or the level of private benefits) is large enough, capital is needed to prevent moral hazard.
that is traded on financial markets, would indirectly reveal the value of \( x_t \) and dispense the regulator from costly auditing\(^{21}\). This idea is explored further in Section 6.

When \( x_t \) is publicly observed, as in our model, the supervisors can have recourse to a second form of market discipline (sometimes called “direct” market discipline) which works by modifying the liability structure of banks. This is the idea behind the "subordinated debt proposal" (Calomiris, 1998, Evanoff and Wall, 2000), which our model allows to analyze formally.

Following this proposal, we assume that banks are required to issue a certain volume \( s \) of subordinated debt, renewed with a certain frequency \( m \). Both \( s \) and \( m \) are policy variables of the regulator. To facilitate comparison with the previous section, we keep constant the total volume of outside finance.\(^{22}\) Thus the volume of insured deposits becomes \( d = 1 - s \). To simplify the analysis, and obtain simpler formulas than Leland and Toft (1996), we assume (as in Ericsson, 2000) that subordinated debt has an infinite maturity, but is renewed according to a Poisson process of intensity \( m \). The average time to maturity of subordinated debt is thus:

\[
\int_0^{+\infty} mte^{-mt}dt = \frac{1}{m}.
\]

In this section, we consider that the regulator can commit to a closure threshold \( x_R \). We focus on the case where \( \lambda x_R < d \), so that deposits are risky, while sub-debt holders (and stockholders) are expropriated in case of closure. We use the same notation as before (for any technology choice \( k = B,G \)):

\[
\begin{align*}
V_k & = \text{value of the bank’s assets}, \\
D_k & = \text{PV of insured deposits}, \\
E_k & = \text{value of equity}, \\
S_k & = \text{market value of subordinated debt}.
\end{align*}
\]

while \( S_k \) denotes the market value of subordinated debt.

Starting with the case where the bank monitors its assets \((k = G)\), the values of \( V_G \) and \( D_G \) are given by simple adaptations of our previous formulas:

\[
\begin{align*}
V_G(x) & = \nu_G x - b + \left[ b - (\nu_G - \lambda)x_R \right] \left( \frac{x}{x_R} \right)^{1-a_G}, \\
D_G(x) & = d - (d - \lambda x_R) \left( \frac{x}{x_R} \right)^{1-a_G}.
\end{align*}
\]

\(^{21}\)Of course, if the bank’s equity is already traded, then this advantage disappears and the question becomes more technical: which security prices reveal more information about banks’ asset value?\(^{22}\) It would be more natural to endogenize the level of outside finance but this would introduce a second state variable and prevent closed form solutions. We are currently working on an extension of this paper in this direction.
\( S_G \) is more difficult to determine. It is the solution of the following Partial Differential Equation, which takes into account the fact that, with instantaneous probability \( m \), subordinated debt is repaid at face value \( s \) but has to be refinanced at price \( S_G(x) \):

\[
\begin{aligned}
    rS_G(x) &= sr + m(s - S_G(x)) + \mu_G x S_G'(x) + \frac{1}{2} \sigma_G^2 x^2 S_G''(x) \\
    S_G(x_R) &= 0,
\end{aligned}
\]

leading to:

\[
S_G(x) = s \left[ 1 - \left( \frac{x}{x_R} \right)^{1-a_G(m)} \right], \quad (17)
\]

where

\[
a_G(m) = \frac{1}{2} + \frac{\mu_G}{\sigma_G^2} + \sqrt{\left( \frac{\mu_G}{\sigma_G^2} - \frac{1}{2} \right)^2 + 2 \frac{r + m}{\sigma_G^2}}. \quad (18)
\]

We immediately notice a first effect of direct market discipline: the exponent \( 1-a_G(m) \) decreases when \( m \) increases. Thus the value of \( S_G \) increases in \( m \). The value of equity becomes:

\[
E_G(x) = V_G(x) - D_G(x) - S_G(x) \\
= \nu_G x - 1 - b + [d + b - \nu_G x_R] \left( \frac{x}{x_R} \right)^{1-a_G} + s \left( \frac{x}{x_R} \right)^{1-a_G(m)}. \quad (19)
\]

When \( m = 0 \), we obtain the same formula as in the previous section (no market discipline): this is due to our convention to keep constant the total volume of outside finance \((s + d = 1)\).

A simple adaptation of formula (19) gives \( E_B \), the value of equity when the bank shirks:

\[
E_B(x) = \nu_B x - 1 + [d - \nu_B x_R] \left( \frac{x}{x_R} \right)^{1-a_B} + s \left( \frac{x}{x_R} \right)^{1-a_B(m)}, \quad (20)
\]

where

\[
a_B(m) = \frac{1}{2} + \frac{\mu_B}{\sigma_B^2} + \sqrt{\left( \frac{\mu_B}{\sigma_B^2} - \frac{1}{2} \right)^2 + 2 \frac{r + m}{\sigma_B^2}}. \quad (21)
\]

Thus a necessary condition for shirking to be eliminated is: \( \Delta \geq 0 \), where

\[
\Delta = x_R[E_G'(x_R) - E_B'(x_R)].
\]
A simple computation gives:

\[
x_R E'_G(x_R) = a_G \nu_G x_R - (a_G - 1)(d + b) - s[a_G(m) - 1],
\]

\[
x_R E'_B(x_R) = a_B \nu_B x_R - (a_B - 1)d - s[a_B(m) - 1].
\]

Thus

\[
\Delta = [a_G \nu_G - a_B \nu_B] x_R - [(a_G - a_B)d + (a_G - 1)b + s\{a_G(m) - a_B(m)\}].
\] (22)

Now, let define \(x_R(m)\) the minimum value \(x_R\) that satisfies inequality \(\Delta(x_R) \geq 0\). Remark that \(x_R(m)\) is implicitly defined by the equation \(\Delta(x_R(m)) = 0\). We show in Appendix C that \((E_G - E_B)(x) \geq 0\) for all \(x \geq x_R(m)\). We deduce

**Proposition 3**: With compulsory subordinated debt, the minimum solvency ratio that prevents bank shirking becomes

\[
x_R(m) = \frac{(a_G - 1)b + (a_G - a_B)d + (a_G(m) - a_B(m))s}{a_G \nu_G - a_B \nu_B}.
\]

or equivalently:

\[
x_R(m) = x_R(0) + s\frac{(a_G(m) - a_B(m)) - (a_G - a_B)}{a_G \nu_G - a_B \nu_B}. \quad (23)
\]

- When \(\Delta \sigma^2 > 0\) this is a U-shaped function of \(m\), with a minimum in \(m^*\).
- When \(\Delta \sigma^2 = 0\) (pure effort problem) \(m^* = +\infty\), which means that \(x_R(m)\) is decreasing for all \(m\).
- Market discipline reduces the need for regulatory bank closures when \(m\) and \(\Delta \sigma^2\) are small.

In order to understand the intuition behind this result, let us recall that \(x_R(m)\) is defined implicitly by the tangency point between the values of equity under the good and the bad technology:

\[
\frac{\partial (E_G - E_B)}{\partial x}(x_R(m), m) = 0.
\]

Given that the value of the bank’s assets and the value of deposits are fixed (once \(x_R\) has been fixed), the changes in the value of equity come from changes in the value of subordinated debt. The question is therefore under what conditions does an increase in the frequency of renewal of subordinated debt increase the derivative of \(S_G\) less than the derivative of \(S_B\) (so that shirking becomes more costly for bankers). Proposition 3 shows that this is true essentially when \(\Delta \sigma^2\) and \(m\) are small.
Figure 4: The impact of an increase in \( m \) on the gains of equityholders when adopting the good technology.\(^{23}\)

Figure 4 illustrates a case where the gap between \( E_G \) and \( E_B \) has been widened by an increase in \( m \) (see dash line), giving the shareholders more incentives to choose the “good” technology. This last property is expressed by the following condition:

\[
\frac{\partial^2 (E_G - E_B)}{\partial x \partial m}(x_R(m), m) > 0.
\]

Proposition 3 clarifies the conditions under which market discipline is a useful complement to solvency regulations: \( \Delta \sigma^2 \) and \( m \) have to be small. Indeed, when \( \Delta \sigma^2 \) is large, \( m^* \) is negative and \( x_R(m) > x_R(0) \) for all relevant values of \( m \) (i.e. positive). In this case the introduction of subordinated debt is counterproductive, since it forces the regulator to increase the minimum capital requirement. The intuition is that the incentives for gambling for resurrection are, in this case, increased by the presence of subordinated debt. On the contrary when \( \Delta \sigma^2 \) is small, \( m^* \) is positive and for \( m \) smaller than \( m^* \), the opposite inequality is true: \( x_R(m) < x_R(0) \). Thus when the risk shifting problem is not too big (\( \Delta \sigma^2 \) small), then for \( m \) small, market discipline reduces the level of regulatory capital.\(^{24}\) We now study how the efficiency of market discipline is affected by the attitude of supervisory authorities.

\(^{23}\)The continuous line represents the difference between \( E_G \) and \( E_B \) as functions of \( x \). The dash line represents the same function after \( m \) has been increased and \( x_R \) modified accordingly.

\(^{24}\)More precisely, the regulatory requirement on equity (\textbf{Tier 1}) is reduced. However, it is easy to see that the total requirement (\textbf{Tier 1} + \textbf{Tier 2}), which includes subordinated debt, is increased.
6 Supervisory action

The “second pillar” of Basel 2 is supervisory review. The Basel Committee states several principles for a sound supervisory review, including: “Supervisors should review and evaluate banks’ internal capital adequacy assessments... [and] take appropriate supervisory action if they are not satisfied with the results of this process” (Basel Committee 2001, p31). However, this is more quickly said than done. Indeed, banking authorities are very often subject to political pressure for supporting banks in distress. In our model, this means providing public funds to the banks who hit the threshold $x_R$. Given our irreversibility assumption, it is indeed always suboptimal (even ex-post) to let banks go below this threshold. On the other hand, closure can be (ex-post) dominated by continuation, when net fiscal costs are not too high. Therefore, whenever a bank hits the boundary $x = x_R$, the government considers the possibility of recapitalizing the bank up to $x_R + \Delta x$ with public funds.

We denote by $\gamma > 0$ the net welfare cost of these public funds, due to the distortions created by the imperfections of the fiscal system. Whenever the government intervenes, the level of recapitalization $\Delta x$ and the new assets value $^{26}$ of the bank $^{27}$ $V_{k,BO}$ (for a technology $k \in \{G, B\}$) are determined by

$$V_{k,BO}(x_R) = \max_{\Delta x \geq 0} \{V_{k,BO}(x_R + \Delta x) - \gamma \Delta x\}. \tag{24}$$

The function $V_{k,BO}$ is determined together with (24), by the usual differential equation that expresses absence of arbitrage opportunities and the “no-bubble” condition: $V_{k,BO}(x) \sim \nu_k x$ for $x \to +\infty$. Therefore $V_{k,BO}(x)$ is necessarily of the type

$$V_{k,BO}(x) = \nu_k x - b_k + \theta_k x^{1-a_k}$$

where the constant $\theta_k$ is determined by (24) and $b_k = 0$ when $k = B$ and $b$ when $k = G$. In Appendix B, we show that the optimal $\Delta x$ is actually $0^+$. This means that the government injects the minimum amount needed to stay above the critical shirking level $x_R$. This can be interpreted as liquidity assistance.

Technically, $x_R$ becomes a reflecting barrier (see for example Dixit, 1993) and the boundary condition for $V_{k,BO}$ becomes:

$$V'_{k,BO}(x_R) = \gamma.$$ 

---

$^{25}$“The Committee views supervisory review as a critical complement to minimum capital requirements and market discipline” (Basel Committee 2001, p30)

$^{26}$This new assets value reflects that future government intervention is anticipated, every time the bank hits again the threshold $x = x_R$.

$^{27}$The letters $BO$ refer to the bail-out operation.
The new formula for $V_G$ is:

$$V_{G, BO}(x) = \nu_G x - b - \left( \frac{\gamma - \nu_G}{a_G - 1} \right) x^{1-a_G} x^{a_G} x_R.$$ 

Notice that this liquidity assistance implies that $V_{G, BO}(x_R)$ is now different from zero, given that banks are allowed to continue after they hit $x_R$. Interestingly, the impact on the value of bank equity, and thus on the incentives for shirking essentially depends on the behavior of the government towards sub-debt holders when the bank hits $x = x_R$. Let us examine successively the case of full expropriation and the case of complete forbearance.

If subdebt holders (and of course equity holders) are wiped out, the decision to rescue the bank only affects the deposit insurance fund, who becomes residual claimant of the assets value of the bank. On the other hand, the differential equation and boundary conditions that characterize the value of subdebt are the same as before and thus $S_k$ is unchanged. Therefore the differential equations that characterize $E_k$ ($k = G, B$) are also unchanged and so are the functions $E_k$ themselves. Market discipline is thus compatible with public liquidity assistance, provided that subordinated debt holders lose their stake if the bank is rescued.

The situation is different in case of complete forbearance, i.e. if subdebt-holders are fully insured when the bank hits the critical threshold $x_R$: $S_k(x_R) = s$. In this case it is easy to see that subdebt becomes riskless and thus its value is identically equal to $s$. But then the subdebt term vanishes from the differential equation that characterizes $E_k$. This means that market discipline becomes totally ineffective.

These results are summarized in the next proposition.

**Proposition 4**: Public liquidity assistance to the banks who hit the critical threshold $x_R$ is compatible with market discipline, provided there is no regulatory forbearance, i.e. if subordinated debt holders are fully expropriated when the bank is rescued. If on the contrary subdebt holders are fully insured, market discipline becomes totally ineffective.

Proposition 4 thus shows that direct market discipline is effective when the credibility of supervisors to close insufficiently capitalized banks is established. We now go further and show how indirect market discipline (i.e. information revealed by the market prices of the securities issued by the banks) can be used to implement a more elaborated regulatory policy, but again when regulatory forbearance is excluded. This illustrates that market discipline is indeed a useful complement to the two other pillars of Basel II: supervision and capital requirements.

---

28 When $\gamma > V_G'(x_R)$, the cost of public funds is so high that the government prefers to close the bank. Whenever $\gamma < V_G'(x_R)$, $V_{G, BO}(x_R) > 0$, which guarantees that bailout is socially preferable (ex-post) to closure.
As already discussed in Section 5, a branch of the academic literature has studied the optimal mix between capital requirements and regulatory audits. In this literature, the value of the bank’s assets (or cash flows $x$ as in our model) is privately known to the banker. It can only be observed by the regulator if a costly audit is performed, which is modeled by a Poisson process, the intensity of which is chosen by the regulator. Since the regulator wants to avoid excessive continuation, he has to set a higher closure threshold than if $x$ was publicly observable. However this threshold can be reduced by an increase in the intensity of auditing, thus suggesting substitutability between supervision and capital requirements.

Our model allows to extend this literature by integrating the 3rd pillar of Basel II, namely market discipline, in this picture. Suppose indeed that the bank has issued at least one security (equity, certificate of deposits or subordinated debt) that is publicly traded on a secondary market. In our model, the price of such a security is a one to one function of the state variable $x$. By inverting this function, the regulator can infer the value of the bank’s cash flows to condition its intervention policy. In such a context, the role of bank supervisors has to be re-examined: instead of a constant intensity of audit across all banks, bank supervisors can adopt a gradual intervention policy (in the spirit of the US regulatory reform following the FDIC Improvement Act). For example, the regulator can set two thresholds $x_R$ and $x_I$ (with $x_R < x_I$), where $x_R$ is as before a closure threshold, but $x_I$ is only an inspection threshold: whenever $x < x_I$, the bank is inspected, the technology chosen by the bank is revealed, and it is closed if and only if the bank has chosen the bad technology ($k = B$).

With this regulatory policy, the value of equity when the technology is “good” is the same as in Section 3 (with $x_R$ replacing $x_L$):

$$E_G(x) = \nu_G x - b - 1 + k_G x^{1-a_G},$$

with

$$k_G = \left[ b + 1 - \nu_G x_R \right] x_R^{a_G-1}.$$  

29Here also, the regulator faces a credibility problem.
30An important empirical literature discusses the predictive power of subordinated debt prices or spreads on banks’ probabilities of failure: see for example Evanoff and Wall (2001), Hancock and Kwast (2001) and Sironi (2001). Covitz et al. (2002) argue that endogeneity of liquidity premia can significantly decrease this predictive power. They argue the subdebt proposal mandating a regular issuance of subdebt can reduce this endogeneity, thus improving information content of spreads and ultimately increase market discipline. Evanoff and Wall (2003) show how subdebt spreads can be used as a complement to capital requirements as a way to reduce regulatory forbearance.
31We assume that the regulatory audit policy is deterministic. It would be easy to consider the more general case of stochastic audits.
However the value of equity when the bank shirks is now a function of $x_I$:

$$E_B(x) = \nu_B x - 1 + k_B x^{1-a_B} \text{ for } x \geq x_I, \ 0 \text{ otherwise},$$

with

$$k_B = [1 - \nu_B x_I]^{a_B-1}.$$

For a given value of $x_I$ (and thus for a given function $E_B$) there is a minimum value of $x_R$ such that $E_G$ remains above $E_B$. It is obtained when the two curves are tangent, as suggested by the figure below.

Figure 5: Optimal capital requirement\textsuperscript{32} $x_R$ for a given inspection threshold $x_I$.

Figure 5 illustrates well how the 3 pillars of Basel II can be optimally mixed: indirect market discipline (i.e. information given to supervisors by market prices of securities issued by banks) could allow a very “light” supervisory policy: only banks with $x \in ]x_R, x_I[$ are audited, and the capital requirement is reduced to $x_R$ (which is less than the one obtained in Proposition 2, in the absence of audit).

7 Concluding Remarks

Our objective in this article is to design a simple dynamic model of commercial bank behavior, where the articulation between the three pillars of Basel II can be analyzed. We

\textsuperscript{32}It can be proven that this minimum level of $x_R$ decreases with $x_I$, which establishes that the substitutability between capital requirements and supervision is maintained when market discipline is introduced.
interpret the first pillar (capital adequacy requirement) as a closure threshold rather than an indirect mean of influencing banks’ asset allocation. We show that market discipline (the third pillar) can be used to reduce this closure threshold, especially if there is a risk of regulatory forbearance. We also re-examine the traditional view on the supervisory role (second pillar).

We suggest that supervisors can modulate the intensity of their interventions (from a simple audit to the closure of the bank) according to reliable signals given by market prices of the securities issued by banks (indirect market discipline). However, two caveats are in order: direct market discipline can only be effective if banking supervisors are protected from political interference, and indirect market discipline cannot be used under all circumstances, since market prices become erratic during crises periods. This shows that Basel II proposals seem to be dangerously insufficient: important reforms of the supervisory system have to be implemented as well, in order to guarantee independence of banking supervisors from political powers and simultaneously proper behavior during future crises.
REFERENCES


Dangl, T. and A. Lehar (2001), “Basle Accord vs Value-at-Risk regulation in Banking”, discussion paper, Department of Business Studies, University of Vienna, Vienna, Austria.


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


Appendix A: proof of Proposition 2

Since by construction $E_G(x_R) = E_B(x_R) = 0$, a necessary condition for elimination of shirking is $E'_G(x_R) \geq E'_B(x_R)$. The minimum value of $x_R$ that satisfies this inequality is defined implicitly by $\Delta(x_R) = 0$ where

$$
\Delta(x_R) = x_R[E'_G(x_R) - E'_B(x_R)].
$$

But formulas (12) and (13) (with $x_R$ replacing $x_L$) imply

$$
x_R E'_G(x_R) = \nu_G x_R + (a_G - 1)[\nu_G x_R - 1 - b],
$$
$$
x_R E'_B(x_R) = \nu_B x_R + (a_B - 1)[\nu_B x_R - 1].
$$

Therefore

$$
\Delta = [a_G \nu_G - a_B \nu_B] x_R - [(a_G - 1)b + a_G - a_B],
$$

which establishes formula (15). Conversely, we have to prove that $(E_G - E_B)(x) \geq 0$ for $x \geq x_R$ when $x_R$ is given by formula (15). The proof is a particular case of the one developed in Appendix C with $m = 0$, $d = 1$ and $s = 0$. Thus we have established that regulation is needed when $x_R > x_L$, which is equivalent to condition (16).

Appendix B: Optimal recapitalization by public funds is infinitesimal (liquidity assistance)

Consider a bank that hits the regulatory threshold $x_R$. When government injects funds $\Delta x$, the new continuation value of the bank at $x_R$ becomes:

$$
V_{k,BO}(x_R) = \max_{\Delta x \geq 0} \{V_{k,BO}(x_R + \Delta x) - \gamma \Delta x\}. \quad (25)
$$

Recall the expression of $V_{k,BO}$:

$$
V_{k,BO}(x) = \nu_k x - b_k + \theta_k x^{1-a_k}.
$$

We want to establish that a finite $\Delta x > 0$ is never optimal. Let us assume by contradiction that such a $\Delta x$ is optimal. The First Order Condition leads to:

$$
V'_{k,BO}(x_R + \Delta x) = \gamma.
$$

Now, because of equation (25),

$$
V_{k,BO}(x_R + \Delta x) = -\gamma \Delta x + V_{k,BO}(x_R)
$$
$$
= V_{k,BO}(x_R) - V'_{k,BO}(x_R + \Delta x) \Delta x
$$
Given that for $\theta_k \neq 0$, $V_{k,BO}$ is either convex or concave, the above equality is only possible when $\theta_k = 0$, $V_{k,BO}$ is affine and the optimal $\Delta x$ is either 0 or infinity. Since by assumption $\gamma > \nu_G$, the optimal $\Delta x$ is always zero.

Appendix C: proof of Proposition 3

We have to prove that

$$(E_G - E_B)(x) \geq 0 \text{ for all } x \geq x_R(m).$$

This result is a straightforward consequence of the lemma:

Lemma 1

$$E'_G(x) > E'_B(x) \text{ for all } x > x_R(m). \quad (27)$$

Proof of Lemma 1:
For all $x > x_R(m)$:

$$xE'_G(x) - xE'_B(x) = (\nu_G - \nu_B)x$$

$$+ (1 - a_G)\{d + b - \nu_G x_R(m)\} \left( \frac{x}{x_R(m)} \right)^{1-a_G}$$

$$- (1 - a_B)\{d - \nu_B x_R(m)\} \left( \frac{x}{x_R(m)} \right)^{1-a_B}$$

$$+ s(1 - a_G(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_G(m)} - s(1 - a_B(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_B(m)}$$

Introducing in this equality, the inequalities $x > x_R(m)$, $1 - a_G < 1 - a_B < 0$, and $(d + b - \nu_G x_R(m)) > 0$, we obtain:

$$xE'_G(x) - xE'_B(x) > (\nu_G - \nu_B) x_R(m) + (1 - a_G)\{d + b - \nu_G x_R(m)\}$$

$$- (1 - a_B)\{d - \nu_B x_R(m)\} \left( \frac{x}{x_R(m)} \right)^{1-a_G}$$

$$+ s(1 - a_G(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_G(m)} - s(1 - a_B(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_B(m)}$$

$$= (a_G(m) - a_B(m)) s \left( \frac{x}{x_R(m)} \right)^{1-a_G}$$

$$+ s(1 - a_G(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_G(m)} - s(1 - a_B(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_B(m)}.$$
Since $a_B(m) < a_G(m)$, this is greater than

$$(a_G(m) - a_B(m))s \left( \frac{x}{x_R(m)} \right)^{1-a_G}$$

$$+ s(1 - a_G(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_G(m)} - s(1 - a_B(m)) \left( \frac{x}{x_R(m)} \right)^{1-a_G(m)}$$

$$= (a_G(m) - a_B(m))s \left( \left( \frac{x}{x_R(m)} \right)^{1-a_G} - \left( \frac{x}{x_R(m)} \right)^{1-a_G(m)} \right)$$

$$> 0 \text{ since } 1 - a_G(m) < 1 - a_G < 0 \text{ and } x > x_R(m).$$

Thus our result.