Optimal Credit Risk Transfer, Monitored Finance, and Banks∗

Gabriella Chiesa†

Abstract

We examine the implications of optimal credit risk transfer (CRT) for bank-loan monitoring, and the incentives for banks to engage in optimal CRT. In our model, properly designed CRT instruments allow banks to insure themselves against loan losses precisely in those states that signal monitoring. We find that optimal CRT enhances loan monitoring and expands financial intermediation, in contrast to the findings of the previous literature. Optimal CRT instruments are based on loan portfolios rather than individual loans and have credit-enhancement guarantees, pretty much as banks do in practice. But the extent of credit enhancement needs to be precisely delimited. Above that exact level, monitoring incentives are undermined (loan quality deteriorates) and wealth is transferred from the bank’s financiers to the bank. Properly designed risk-based capital requirements are shown to prevent such a wealth transfer and to provide banks with the incentive to engage in optimal CRT.

Key Words: Credit Risk Transfer, Monitoring Incentives, Prudential Regulation; JEL classification: G21, D82, G28, D61.

∗I thank Vincenzo Denicolo’, Alan Morrison, Oren Sussman, Elu Von Thadden, Anjan Thakor, and Janping Qi for their comments and suggestions. I also benefited from comments by seminar participants at Venice, Salerno, Tilburg, the Ente Einaudi Institute for Economics and Finance (Rome), the FIRS Conference on Banking, Corporate Finance and Intermediation in Shanghai, the European Finance Association Meeting in Zurich, and the BIS-CEPR-JFI Conference on Risk Management and Regulation in Banking in Basel.
†Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy; gabriella.chiesa@unibo.it
1 Introduction

Traditionally, the banking paradigm adhered quite well to real-world practice: loans were held on the balance sheet until maturity or default, and the bank’s main risk-management tool was portfolio diversification. Retaining the loans and diversifying away their idiosyncratic risks allowed debt-financed banks to retain monitoring incentives and hence perform the role of delegated monitors (Diamond, 1984; Ramakrishnan and Thakor, 1984).

While banks have continued to raise funds mainly via debt (including deposits), they now engage extensively in loan sales and more generally in credit risk transfer (CRT), as documented by BBA (2004), ECB (2004), BIS (2005a), Minton et al (2005), Duffie (2007). The transfer of risk on unrated borrowers, who typically need bank monitoring, is increasing steadily (Fitch Ratings, 2004). This has attracted the attention of policy makers and national and supranational supervisors, prompting an enormous number of reports (e.g. BBA, 2004; BIS, 2003; BIS, 2005a; BIS, 2005b; IMF, 2002). The ambivalence of these reports about the merits of CRT is well captured by Warren Buffett (2002), who argues that CRT may harm the stability of the financial sector. By contrast, Alan Greenspan (2005) highlights the evidence that in the early 2000s US recession corporate failures neither caused banking failures nor harmed the financial sector as a whole, owing to the widespread use of loan securitization and credit derivatives. That is, Greenspan notes the merits of CRT as a risk-management tool, whereas Buffett points out the potential defect of allowing banks to forfeit their monitoring/screening role. Current credit market turmoil, also triggered by the asset quality problems of the so-called sub-prime market, seems to support Buffett’s view and raise doubts about the provision of incentives underlying the new pattern of intermediation, the so-called originate-to-distribute (OTD) model. But, is the OTD model necessarily harmful, or are the CRT instruments used that distort incentives. If so, why are they used? What is the role for prudential regulation?

The literature generally finds that CRT weakens banks’ monitoring incentives, and hence undermines financial stability.1 This paper revisits the

---

1The theoretical literature on financial intermediation stresses banks’ monitoring role (Campbell and Kracaw, 1980; Diamond, 1984; Fama, 1985; Hellwig, 1991; Bhattacharya and Chiesa, 1995; Holmstrom and Tirole, 1997; see also the banking literature review by Bhattacharya and Thakor, 1993). There is also ample evidence that bank monitoring
issue, imposing no exogenous restrictions on the nature of CRT instruments. In keeping with the evidence that correlation of defaults is driven by the business cycle (BIS, 2005b; Keenan, 2000), we allow for aggregate risk: loans are subject to idiosyncratic risks and to a common, macroeconomic, risk factor. As in many other models, monitoring by banks improves the expected return from loans. However, in our model the effect of such monitoring is most valuable in economic downturns; monitoring lowers loan default probability, but still results in some defaults. Properly devised CRT instruments allow banks to insure themselves against losses precisely in those states that signal monitoring. We find that optimal CRT enhances loan monitoring and expands financial intermediation, in contrast to the findings of the previous literature. We also show that optimal CRT instruments are based on loan portfolios rather than individual loans, and have credit-enhancement guarantees. In our model, optimal CRT maximizes the lending level for which the bank finds it incentive-compatible to monitor (its incentive-based lending capacity) for any given amount of capital. Or, equivalently, optimal CRT allows banks to economize on capital: any given amount of incentive-compatible lending can be sustained with minimum capital.

Portfolio CRT instruments (e.g., loan portfolio securitization with credit-enhancement guarantees) are pretty much the instruments used by banks. But we find that the extent of credit enhancement needs to be precisely delimited; it must be within a defined interval. Outside that interval, monitoring incentives are undermined. The bank will never provide too little credit enhancement; the dilution of monitoring incentives (the deterioration of loan quality) would be factored into the pricing of the insufficiently-credit-enhanced securities. However, the bank has an incentive to provide excessive credit enhancement. This undermines monitoring incentives and transfers


3 This fits the empirical evidence that banks that engage in CRT have greater leverage and make more business loans (Cebenoyan and Strahan, 2004).

4 The BBA credit derivatives reports (2001/02, 2003/04) highlight the rapid increase in portfolio products and among them in loan portfolio securitization, which is estimated to account for about 26% of products by the end of 2004.
wealth from depositors to the bank in a way that makes the bank better off. We show that properly designed risk-based capital requirements prevent such a wealth transfer and provide banks with the incentive to engage in optimal CRT.

The intuition for our results is as follows. By monitoring or screening loan applicants, the bank raises the quality of its borrowers’ projects, which improves loan portfolio performance in bad times. But the bank is subject to moral hazard even when it is perfectly diversified and all idiosyncratic risk is diversified away. It may be tempted to bet that fortunate external circumstances (a macroeconomic upturn) will sustain the borrowers’ performance, avoid costly monitoring and shift the (unmonitored) loan losses that emerge in downturns to its financiers. The bank’s moral hazard problem is addressed by two instruments: reward for monitoring (the carrot), and capital per unit of lending (the stick), so that the more likely losses due to poor monitoring are borne partially by the bank’s owners (insiders). Since bank capital is a scarce resource, an optimal contract for the bank to raise funds minimizes the capital injection (stick) and stresses the reward (carrot). It uses the information conveyed by the bank’s loan portfolio return and rewards the bank as much as possible for those outcomes that signal monitoring, in the sense of Holmstrom (1979). Our point, however, is that these monitoring-revealing outcomes are not the "high" outcomes in terms of the returns on the bank’s loan portfolio. In our model, high outcomes may result from good luck (an economic upturn) rather than monitoring. As a result, debt financing for banks is suboptimal. With debt, the better the outcome, the greater the debtor’s income: the bank is rewarded for good luck rather than for monitoring.

Optimal CRT addresses the shortcomings of deposit/debt financing. It reallocates the bank’s income from lucky states, when its portfolio returns are high, to those that are more informative about its monitoring effort. This decreases the amount of capital per unit of lending the bank must inject to find it incentive-compatible to monitor, and incentive-based lending capacity expands. It raises more outside funds and its lending expands,

---

5 See Holmstrom and Tirole (1997) for a discussion of costs of bank capital. For reasons of adverse selection about the quality of bank loan portfolios, not modelled here, many scholars, e.g., Bolton and Freixas (2000), hold that bank capital is more costly than other sources of bank funding.
but it still monitors. This explains why CRT increases efficiency. Optimal CRT is based on loan portfolio and is backed by a precise extent of credit-enhancement guarantees. The bank is insured against loan losses precisely in those states that signal monitoring; its incentive-based lending capacity is maximized.

One result we get, then, is that contrary to what is often claimed CRT enhances banks’ monitoring role, if used properly. Much of the previous literature argues that CRT necessarily weakens monitoring incentives (Pennacchi, 1988; Gorton and Pennacchi, 1995; Duffee and Zhou, 2001; Morrison, 2005; Behr and Lee, 2005; Parlour and Plantin, 2007; Parlour and Winton, 2007). In all these works debt financing maximizes the reward for monitoring (see Section 4). Additional arrangements, such as CRT, can only reduce monitoring incentives. In contrast, in our paper, optimal CRT creates value on incentive-based grounds: it lowers the amount of capital a bank must have at stake in order for a monitoring incentive to subsist. It thus expands incentive-based lending capacity, and hence the availability of monitored finance. Furthermore, as a risk-management tool, CRT complements diversification: the prior construction of diversified loan portfolios enables optimal CRT to insulate the bank from exogenous risk (given monitoring). Monitoring incentives are enhanced.

However, for these value/welfare gains to obtain, banks must have the incentive to engage in optimal CRT. The bank, having promised to engage in optimal CRT, may not honor this promise once having raised funds – the time-inconsistency/commitment problem which arises in hedging (Smith and Stulz, 1985; see Section 7). And indeed, after having borrowed funds and made loans, an unregulated bank would have an incentive to retain loan risks. We examine the role of market forces and of prudential regulation in solving this problem (Section 6).

The rest of this paper is organized as follows. Section 2 presents the model. Section 3 derives the optimal bank financing contract. Section 4 analyses banks’ monitoring incentive problem under debt financing. Section 5 shows that debt financing with CRT implements the optimal contract, provided banks can commit to engage in optimal CRT after borrowing. Section 6 analyses banks’ ex-post incentives to engage in optimal CRT, and the role for prudential regulation. In Section 7 we discuss related literature on CRT in greater detail. Section 8 concludes.
2 The Model

A bank has the opportunity to finance a portfolio (continuum) of loans whose size \( L \) is derived endogenously. It funds lending out of internal funds (capital) and outside finance from final investors. The supply of outside finance is perfectly elastic at a gross rate of return that is normalized to one (i.e. the risk-free net interest rate is zero). That is, final investors are assumed to make zero profits. The bank acts on behalf of its shareholders (insiders), whose equity holdings constitute the bank’s endowment of inside capital, \( K \).

2.1 Project Technology and Monitoring

Lending consists in project financing. A project requires one unit of resources at date 0 and gives the bank a return \( X \in \{0, R\} \) at date 1. The probability of success, \( \Pr(X = R) \), depends on the project type \( t \in \{g, b\} \) and on the realization at date 1 of a common, macroeconomic, risk factor \( \theta \in \{\overline{\theta}, \underline{\theta}\} \). Here \( \overline{\theta} \) denotes the good realization, an economic upturn, and occurs with probability \( p > 0 \); \( \underline{\theta} \) denotes the bad realization, the downturn, and occurs with probability, \( 1 - p \). If at date 1 the common-factor realization is \( \overline{\theta} \) (upturn), then a project succeeds for sure regardless of its type. If the common-factor realization is \( \underline{\theta} \) (downturn), then a type \( g \) project succeeds with probability \( \alpha < 1 \) and a type \( b \) project succeeds with probability \( \underline{\alpha} < \alpha \).

A project is of type \( g \) if and only if it has been monitored by the bank at the beginning of the period. Table 1 shows the distribution of project return conditional on the bank’s action and the realization of the common risk factor.

TABLE 1 ABOUT HERE

Monitoring may consist in project financing. A project requires one unit of resources at date 0 and gives the bank a return \( X \in \{0, R\} \) at date 1. The probability of success, \( \Pr(X = R) \), depends on the project type \( t \in \{g, b\} \) and on the realization at date 1 of a common, macroeconomic, risk factor \( \theta \in \{\overline{\theta}, \underline{\theta}\} \). Here \( \overline{\theta} \) denotes the good realization, an economic upturn, and occurs with probability \( p > 0 \); \( \underline{\theta} \) denotes the bad realization, the downturn, and occurs with probability, \( 1 - p \). If at date 1 the common-factor realization is \( \overline{\theta} \) (upturn), then a project succeeds for sure regardless of its type. If the common-factor realization is \( \underline{\theta} \) (downturn), then a type \( g \) project succeeds with probability \( \alpha < 1 \) and a type \( b \) project succeeds with probability \( \underline{\alpha} < \alpha \).

A project is of type \( g \) if and only if it has been monitored by the bank at the beginning of the period. Table 1 shows the distribution of project return conditional on the bank’s action and the realization of the common risk factor.

TABLE 1 ABOUT HERE

Monitoring may consist either of services tailored to the borrower, or a set of constraints on his choice among projects by appropriate debt covenants, whose fulfillment is then monitored. For example, an unconstrained borrower might select the specific project for possible private benefits. Monitoring may also consist of screening borrowers, i.e. costly testing of the
borrower’s creditworthiness in an adverse-selection environment.\(^6\)

Monitoring has a cost to the bank of \(F > 0\) per project. This is a non-pecuniary effort cost. \(F\) may also be interpreted as the opportunity cost of eschewing insider lending, forgoing the private benefits of collusion with the borrower on a bad (type b) project.\(^7\)

We further specify the model by making the following assumptions:

Assumption A1. An unmonitored project has negative net present value:

\[
pR + (1 - p)\alpha R < 1 \quad ,
\]

(A1)

Assumption A2. A monitored project has positive net present value:

\[
pR + (1 - p)\alpha R > 1 + F \quad ,
\]

(A2)

Assumption A3. The bank’s monitoring/screening choices are unobservable.

We assume that the bank either monitors the entire portfolio or does not monitor at all. That is, we rule out monitoring of only some loans. This simplifies the exposition with no loss of generality. Indeed, since the monitoring technology has constant return to scale, partial monitoring is never optimal.

2.2 Portfolio Outcomes

For simplicity, loan portfolios are assumed to be perfectly diversified, so that idiosyncratic risk is diversified away. Outcomes are thus centered on the mean, which depends on the realization of the common-factor and on the bank’s monitoring choice.

Let \(s\) denote the loan-solvency rate realization. For a diversified portfolio, \(s\) can be equal to 1, \(\alpha\) or \(\bar{\alpha}\). The outcome \(s = 1\) occurs with the probability \(p\) that the common factor realization is \(\bar{\theta}\) (upturn), hence all loans perform. With probability \(1 - p\), the common factor realization is \(\tilde{\theta}\) (downturn), and the outcome is \(s = \alpha\) if the bank has monitored and \(s = \bar{\alpha}\) otherwise.

\(^6\)In this case, the bank’s monitoring cost is the cost of the test, divided by the probability that the borrower has a good (type g) project.

\(^7\)See Repullo and Suarez (1998) for an explicit analysis of collusion between the entrepreneur and the monitoring/informed lender.
Clearly, certain outcomes signal the bank’s action. For example, the outcome \( s = \alpha \) perfectly reveals that the bank has not monitored. Likewise, \( s = \alpha \) reveals that the bank has monitored, in spite of \( \alpha \) not being the highest portfolio outcome. The highest outcome of course, is \( s = 1 \), but this is just good luck. As we shall see later (Section 4), in this framework debt financing becomes suboptimal.

3 Optimal Contract

What is the optimal contract with outside finance providers, i.e. the one that maximizes bank’s profits? Since unmonitored projects have negative net present value and monitored projects positive, an optimal contract necessarily implies the incentive to monitor. Moreover, final investors make zero profits. It follows that the bank’s profit per unit of lending is the surplus generated by a monitored loan, \( pR + (1 - p)\alpha R - (1 + F) \). Clearly this is positive, so the bank’s profits are increasing in lending.

It follows that an optimal contract maximizes the amount of outside finance, and hence bank’s lending, subject to the incentive compatibility constraint (the bank must find it profitable to monitor) and the participation constraint (investors make zero profits). It is also easy to show that the optimal contract must use all available information to reward the bank “as much as possible” for observable outcomes that signal monitoring. To see why, consider contracts of the form \((W_1, W_\alpha, W_\bar{\alpha})\), where \( W_s \) is the bank’s income. Investors’ income is \( sLR - W_s \) with \( s \in \{1, \alpha, \bar{\alpha}\} \). The case \( s = 1 \) results from good luck (the common factor realization is \( \theta = \bar{\theta} \)); thus income \( W_1 \) can be defined as the bank’s reward for good luck. If \( s = \alpha \), then necessarily the bank has monitored; \( W_\alpha \) is the bank’s reward for monitoring. If \( s = \bar{\alpha} \), then necessarily the bank has not monitored; \( W_\bar{\alpha} \) is the bank’s reward for not monitoring. Clearly, the optimal contract has \( W_\bar{\alpha} = 0 \); that is, the bank must be penalized as much as possible for not monitoring. In principle, also \( W_1 \) should be as small as possible. However, if \( W_1 < W_\alpha \) the contract would be vulnerable to portfolio outcome falsification: in the upturn the bank would profit by destroying part of the lending income (as by forgiving some of borrower debt) so as to mimic the performance of a monitored portfolio in the downturn, thereby getting \( W_\alpha \) instead of
Thus, the bank’s income schedule needs to satisfy the monotonicity constraint $W_1 \geq W_\alpha$.

The optimal contracting problem can then be expressed as follows:

$$\max_{L,W_1,W_\alpha} \{pW_1 + (1 - p)W_\alpha - FL - K\}$$

s.t.

$$pW_1 + (1 - p)W_\alpha - FL - K \geq pW_1 - K$$

$$p(RL - W_1) + (1 - p)(\alpha LR - W_\alpha) = L - K$$

$$W_1 \geq W_\alpha$$

The objective function is the bank’s expected profit conditional on monitoring. This equals the expected income, $pW_1 + (1 - p)W_\alpha$, less monitoring costs, $FL$, less internal finance, $K$. Condition (2) is the incentive compatibility constraint: the bank’s profit conditional on monitoring must be at least as great as without monitoring. This is $pW_1 - K$ because $W_\alpha = 0$. Condition (3) is the investors’ zero profit constraint, that the amount of finance investors provide, $L - K$, equals their expected income – the left-hand side of (3). Condition (4) is a monotonicity constraint on the bank’s income schedule discussed above.

Rearranging terms, the problem reduces to:

$$\max_{L,W_1,W_\alpha} L \left[ pR + (1 - p)\alpha R - (1 + F) \right]$$

s.t.

$$L \leq \frac{(1 - p)}{F}W_\alpha$$

$$p(RL - W_1) + (1 - p)(\alpha LR - W_\alpha) = L - K$$

$$W_1 \geq W_\alpha$$

Since $[pR + (1 - p)\alpha R - (1 + F)] > 0$, maximizing profits amounts to maximizing lending $L$. Now suppose that

$$[p + (1 - p)\alpha] R - \frac{F}{1 - p} \geq 1$$
In this case, the bank would be effectively unconstrained and could borrow an infinite amount: if (5) holds, then for any \( L \) there exist \( W_\alpha, W_1 \), such that the incentive-compatibility and the monotonicity constraints hold and investors make non-negative profits (for example, \( W_1 = W_\alpha = \frac{FL}{(1-p)} \)). We rule out this rather uninteresting case by assuming:

\[
pR + (1-p)\alpha R - \frac{F}{1-p} < 1. \tag{A3}
\]

Next, define \( c^* \)

\[
c^* \equiv \{1 - [p + (1-p)\alpha] R\} + \frac{F}{1-p}.
\]

By (A1) – (A2), we have \( (1-p)(\alpha - \alpha) R > F \), and \( [p + (1-p)\alpha] R > (\alpha - \alpha) R \); hence \( c^* < 1 \). We have:

**Proposition 1**  
At the optimum, the bank’s lending capacity is:

\[
L^* = \frac{K}{c^*},
\]

and the bank’s income is:

\[
W^*_\alpha = W^*_1 \equiv L^* \{[p + (1-p)\alpha] R - 1\} + K, \quad W^*_2 = 0.
\]

The proof is very simple: by the incentive compatibility constraint (2’), maximization of \( L \) requires the reward for monitoring, \( W_\alpha \), to be as great as possible. The investor’s participation constraint then implies that \( W_1 \) must be as small as possible. This means that the monotonicity constraint binds: \( W_1 = W_\alpha \). Substituting into the participation constraint gives \( W_1 = W_\alpha = L \{[p + (1-p)\alpha] R - 1\} + K \). The incentive constraint then reduces to \( c^* L \leq K \). Since \( c^* > 0 \) (by \( (A3) \)), the highest lending level that satisfies the incentive constraint – the bank’s lending capacity – is \( L^* = \frac{K}{c^*} \).

Could the bank do better than \( L^* \)? Yes, if it could credibly promise to give final investors a larger portion of the portfolio outcome in the upturn – which is merely due to good luck. The reward for monitoring would then increase, and so therefore would lending and profits. But, this promise is not credible, as it would violate the monotonicity constraint.
For future reference, the bank’s expected profit with the optimal contract is
\[ \Pi^* = pW_1^* + (1 - p)W_\alpha^* - FL^* - K \]
\[ \equiv pW_1^* - K \quad (\text{by } L^* \equiv \frac{K}{\alpha}) . \]

Having thus characterized the optimal contract, in the remainder of the paper we discuss the issue of implementation. In the real world, bank first collect deposits, then make loans, and then possibly engage in CRT. We show that the optimal contract cannot be implemented with debt, but can be implemented with debt financing \textit{cum} optimal CRT.

### 4 Debt Financing

In this section we ask what can be implemented with debt. What is the maximum amount of lending that makes the bank still willing to monitor, when \( L - K \) is raised by debt? Suppose the bank issues debt \( D = L - K \) with a face value such that it pays \( D_\alpha \) if the portfolio-solvency rate is \( s \).

The problem reduces to determining the maximum value of \( L \) for which the incentive compatibility constraint still holds. With debt, the incentive compatibility constraint becomes:
\[ p(LR-D_1)+(1-p)(\alpha LR-D_\alpha)-FL-K \geq p(LR-D_1)+(1-p)\{\alpha LR-D_\alpha\}-K \]
\[ \text{(7)} \]

The left-hand side of (7) is the bank’s expected profit conditional on monitoring; the right-hand side, on not monitoring.

Clearly, the bank must be solvent in state \( s = \alpha \); for otherwise it would not monitor. Then it is necessarily solvent in \( s = 1 \), and therefore the face value of the debt that satisfies the investors’ zero-profit constraint is \( L - K \). It follows that \( D_1 = D_\alpha = L - K \), and the incentive constraint reduces to:

\[ (1 - p) [\alpha LR - (L - K)] - FL \geq (1 - p) \max \{\alpha LR - (L - K), 0\} \]

If the bank is solvent in state \( s = \alpha \), i.e. if \( \alpha LR \geq (L - K) \), then the incentive constraint becomes:
\[ R(\alpha - \alpha) \geq \frac{F}{1 - p} \]

\[ ^{8} \text{If the loan-portfolio revenue, } sLR, \text{ is not lower than the face value of the debt, then } D_\alpha \text{ equals the face value – the bank is solvent. Otherwise, } D_\alpha = sLR. \]
which always holds by (A1)-(A2). This merely says that if it is financially unconstrained the bank would always monitor. Thus, \( L \) will be raised at least up to the point where the bank is financially constrained, i.e. where it is insolvent in state \( s = \alpha \). When \( \alpha LR < (L - K) \), the incentive constraint is:

\[
\alpha LR - (L - K) \geq \frac{FL}{1 - p}
\]

which holds if and only if:

\[
L \leq \frac{K}{(1 - \alpha R) + \frac{F}{1 - p}}.
\]

Define \( c^D \equiv (1 - \alpha R) + \frac{F}{1 - p} \). Thus, we have

**Proposition 2** If the bank is financed with debt only, the optimal contract cannot be implemented. The bank’s lending capacity is:

\[
L^D = \frac{K}{c^D} < L^*,
\]

and its income is:

\[
W^D_1 = L^D (R - 1) + K ; W^D_\alpha = L^D (\alpha R - 1) + K ; W^D_\alpha = 0.
\]

With a debt contract for financing, the better the outcome the greater the bank’s income: \( W^D_1 > W^D_\alpha \). This means that good luck is over-rewarded and monitoring is under-rewarded, \( W^D_\alpha < W^*_\alpha \). As a consequence, the minimum amount of capital per lending unit needed to make it incentive-compatible to monitor increases, \( c^D > c^* \), and the bank’s lending capacity therefore shrinks, \( L^D < L^* \).

In our model debt financing is distortionary because monitoring is most valuable when the macroeconomic shock is adverse (downturn); monitoring removes the possibility of a high proportion of loan defaults, but still results in some defaults. The bank’s payoff should be reallocated from the high state (merely luck) to the monitoring-revealing state, so that the bank gets a higher reward in this state. But debt does exactly the opposite and is
therefore suboptimal.\footnote{Formally, in our model the loan-portfolio return distribution fails to satisfy the Monotone-Likelihood-Ratio Property (MLRP) with respect to monitoring effort. Failure of MLRP is sufficient but not necessary for debt to be suboptimal. In Chiesa (1992), for instance, MLRP holds but debt is suboptimal because it offers poor incentives in bad states.}

In the previous CRT literature (Pennacchi, 1988; Gorton and Pennacchi, 1995; Duffee and Zhou, 2001; Morrison, 2005; Behr and Lee, 2005; Parlour and Plantin, 2007; Parlour and Winton, 2007) the monitoring-revealing outcomes are the "high" outcomes.\footnote{This means that the loan-portfolio return distribution satisfies the Monotone-Likelihood-Ratio Property (MLRP) with respect to monitoring effort; and so debt can be optimal (Innes, 1990).} In our model this corresponds to the special case $p = 0$. If $p = 0$, then the monitoring-revealing state is also the most favorable. In this case, the reward for monitoring under debt financing is the same as in an optimal contract, $W^D_\alpha \equiv W^*_\alpha$. Differently put, when $p = 0$, $c^D \equiv c^* < 0$: when all idiosyncratic risk is diversified away, the first-best optimum is attained, and under an optimal contract the bank’s lending level is unconstrained – as in Diamond (1984) and in Laux (2001), in an industrial organization context. Since debt maximizes the bank’s reward for monitoring, additional arrangements, such as CRT, can only reduce monitoring incentives. This explains the negative effect of credit risk transfer found in the literature. But when debt financing under-rewards monitoring, CRT is the right tool for restoring incentives, as we now show.

5 Optimal Credit Risk Transfer

Let us now suppose that after collecting deposits and making loans, banks engage in CRT. One problem with this strategy is that the bank, having promised to engage in the optimal risk transfer deal, may not honor this promise once having raised funds. To proceed, we first assume that the bank can commit to engage in such a deal. Later we address the issue of commitment (Section 6). Consider the following arrangement:

1. The bank raises deposits $L^* - K$ and finances the optimal loan portfolio $L^*$
2. It then forms a special purpose vehicle and securitizes/sells the entire loan portfolio for a total price $P_0$
3. The bank credit-enhances the deal by giving investors the option to sell their claims back to the bank for a total price $P$. To back this guarantee, the bank injects $P$ as cash collateral.

The above contract defines the CRT mechanism $(P_0, P)$. The total price $P$ at which investors can sell their claims back to the bank defines the protection provided to investors (security buyers). Clearly, it must be $P \leq P_0$ – the cash collateral injection required is feasible.

Define $P \equiv \left[ \alpha R + \frac{E}{1-p} \right] L^*$ and $\overline{P} \equiv \alpha L^*$ as the lower and upper bound of the protection that an optimal CRT mechanism provides.\(^{11}\)

**Proposition 3** Any credit-risk-transfer mechanism with $P^* \in [L^*, \overline{P}]$, and $P_0^* \equiv p L^* R + (1-p) \alpha L^* R$, implements the optimal contract.

**Proof.** Let us consider the bank’s income under the CRT mechanism $(P_0^*, P^*)$ in each possible state $s \in \{1, \alpha, \alpha \}$. If $s = 1$, loan-portfolio revenue $L^* R$ realizes. Investors forgo their options, and the bank collects the cash collateral. Thus its total cash is $L^* [p R + (1-p) \alpha R]$, out of which the bank repays $L^* - K$ to depositors, ending up with income $[p + (1-p) \alpha] L^* R - (L^* - K) \equiv W_1^*$.

If $s = \alpha$ (a state that occurs only if the bank has monitored), loan-portfolio revenue $\alpha L^* R$ realizes. Investors again forgo their options, since $\max(\alpha L^* R, P^*) = \alpha L^* R$ (this follows from $P^* \in [L^*, \overline{P}]$). Then the bank collects the cash collateral and repays depositors, ending up with income $[p + (1-p) \alpha] L^* R - (L^* - K) \equiv W_\alpha^*$.

Finally, if $s = \overline{\alpha}$, loan-portfolio revenue $\alpha L^* R$ realizes. Investors now exercise their options and so the bank loses the cash collateral. Thus, the bank has total cash $[p + (1-p) \alpha] L^* R - P^* + \alpha L^* R$, and its income is

$$\max \{ [p + (1-p) \alpha] L^* R - P^* + \alpha L^* R - (L^* - K), 0 \} .$$

Since $L^* \equiv \frac{K}{\alpha}$, we have $[p + (1-p) \alpha] L^* R - \overline{P} + \alpha L^* R \equiv L^* - K$, and since $P^* \geq \overline{P}$, the bank’s income is nil. Hence, $W_\alpha^* \equiv W_\alpha^*$. This proves that $(P_0^*, P^*)$ implements the optimal contract. $\blacksquare$

Proposition 3 shows that the bank’s income with the securitization scheme $(P_0^*, P^*)$ is the same as in the optimal contract, which maximizes lending

\(^{11}\)By (A1) – (A2), we have $(1-p) (\alpha - \overline{\alpha}) R > F$; hence $\overline{P} > \overline{P}$. 

14
capacity. Depositors and investors break even, and all loans are monitored. Note that the CRT mechanism is based on loan portfolios rather than individual loans, and is backed by an extent of credit enhancement in the interval $[P, \overline{P}]$.

Portfolio securitization, backed by the required extent of credit enhancement, resolves the shortcomings of debt financing: it shifts income from the "lucky" high state to the monitoring-revealing state, so giving the bank a higher reward in the latter. This would not be true for a partial loan sale (as in Gorton and Pennacchi, 1995; Pennacchi, 1988); in a partial sale, or equivalently a sale with partial recourse, the bank is exposed to common exogenous risk: its income in the monitoring-revealing state $s = \alpha$ is lower than that in the high state $s = 1$. That is, the reward for monitoring is not maximized, so the bank's lending capacity falls below the maximum.

Any protection in the interval $[P, \overline{P}]$ supports the monitoring equilibrium. Outside this interval, however, the bank's monitoring incentives are undermined. To be precise, it is clear from the proof of Proposition 3 that if $P < P^*$ then $W_\alpha > W^*_\alpha$, and if $P > \overline{P}$ then $W_\alpha < W^*_\alpha$; in both cases, the bank's incentive-compatibility constraint for monitoring optimal loan portfolio $L^*$ fails to hold.

6 The Bank’s Incentive to Engage in Optimal CRT: Prudential Regulation

As we noted above, banks raise funds and make loans before engaging in CRT. But does the bank still have an incentive to engage in optimal CRT once the funds have been raised? In fact it does not. An unregulated bank would have an incentive to credit enhance the CRT deal excessively, to such an extent that effectively amounts to retaining the entire credit risk.

Consider again the sequence of events detailed in Section 5. Once the bank has borrowed $L^* - K$ and lent $L^*$, does it find it profitable to engage in the optimal CRT deal with $P^* \in [P, \overline{P}]$, or does it make a higher profit by setting $P < P$ or $P > \overline{P}$?

Clearly, if $P < P$, the dilution of monitoring incentives would be factored into the pricing of the insufficiently-credit-enhanced securities. These would
be priced at

\[ P_0 \equiv pL^*R + (1 - p) \max (P, \alpha L^*R) < P^*, \]

because \( P < \bar{P} < \alpha L^*R \), so the bank would be worse off.\(^{12}\)\(^{13}\)

However, consider the deal \((P_0, P)\), with \( P \leq L^*R \) and \( P > \bar{P} \equiv \alpha L^*R \), and

\[ P_0 = pL^*R + (1 - p)P \]

This contract provides investors with excessive protection, \( P > \bar{P} \), but since this is factored into the pricing, investors would break even as they do in the optimal CRT deal. However, the bank would make more profit, while depositors would lose. To see why, consider the bank’s payoff in state \( s \in \{1, \alpha\} \).

In state \( s = 1 \), investors forgo their options and the bank’s income is

\[ W_1 = P_0 - (L^* - K) \equiv W_1^* + (1 - p)(P - \bar{P}) \]

In state \( s = \alpha \), investors now exercise their options and the bank ends up with income

\[ W_\alpha = \max \{(P_0 - P) + \alpha L^*R - (L^* - K), 0\} \equiv \max \{W_\alpha^* - p(P - \bar{P}), 0\} \]

Clearly, an excessively credit-enhanced deal \((P_0, P)\) under-rewards monitoring, \( W_\alpha < W_\alpha^* \), and over-rewards good luck, \( W_1 > W_1^* \). As a result, the

\(^{12}\)That is, \( \Pi < \Pi^* \equiv pW_1^* - K \). To show this, note that

\[ \Pi \equiv p \max [P_0 - (L^* - K), 0] + (1 - p) \max [P_0 - P + \min (P, \alpha L^*R) - (L^* - K), 0] - K \]

is the bank’s profit with an insufficiently enhanced deal, \( P < \bar{P} \). If the term inside curley brackets is positive, that is the bank is solvent in \( s = \alpha \), then the inequality \( \Pi < \Pi^* \) certainly holds, because depositors do not make losses, investors break even, and hence the value loss resulting from not monitoring bears on the bank. If the bank is insolvent in \( s = \alpha \), then the inequality reduces to:

\[ W_1^* > \max [P_0 - (L^* - K), 0], \]

which again holds because \( W_1^* \equiv [p + (1 - p)\alpha] L^*R - (L^* - K) \), and \( [p + (1 - p)\alpha] L^*R > P_0 \equiv pL^*R + (1 - p) \max (P, \alpha L^*R) \) (by \( \alpha L^*R > \bar{P} > P \)).

\(^{13}\)A referee has pointed out that a bank could still cheat by selling loans and investing the money raised in option contracts that pay off in state \( s = 1 \). Clearly, this is feasible only if the option contract is not subject to capital requirements, it then suggests a further argument in favor of capital requirements on retained risks.
bank would not monitor, failing in the downturn. Its profit would then be

$$\Pi = pW_1 - K \equiv p \left[ W_1^* + (1 - p) (P - \overline{P}) \right] - K$$

(8)

which is higher than the profit it makes in the optimal scheme, $$\Pi^*$$, whenever $$P > \overline{P}$$ (compare (8) and (6)). The intuition is that lending $$L^*$$ and retaining risk in excess of the optimal amount (i.e., raising the protection provided to investors above $$\overline{P}$$) is tantamount to lending in excess of incentive-based capacity and shifting losses to depositors. That is, it entails an ex-post wealth transfer from depositors to the bank. The greater the protection, the larger this wealth transfer.

Anticipating that the bank will retain the entire credit risk, depositors will ration the bank so that the lending it can make does not exceed $$L^D$$, the maximum lending volume that will be monitored in the absence of CRT.

This time inconsistency problem could be addressed in various ways. For example, the optimal use of CRT might be sustained by reputational mechanisms (Boot, Greenbaum, and Thakor, 1993), or by (the threat of) bank runs à la Diamond and Rajan (2001), or, more generally, by demandable debt: If the bank refrained from optimal CRT after lending $$L^*$$, funds would be withdrawn and the bank would be unable to refinance its debt (Calomiris and Kahn, 1991). Clearly, the threat of withdrawals/runs needs to be credible; deposits ought to be uninsured. In the reminder of this section we briefly discuss another possible way to restore efficiency, namely, through capital requirements on loans conditioned on the extent of retained risk.

In our setting, the capital requirement must be conditioned on the ratio $$P/L$$, i.e., the protection provided to security buyers per unit of loans. Specifically, suppose that for any given $$\frac{P}{L} \geq \alpha R$$, the capital requirement is

$$c_{P/L} = c^* + p \left( \frac{P}{L} - \alpha R \right).$$

With such a capital requirement in place, if the bank after having made $$L$$ loans engages in the CRT deal with protection $$P$$, then regulatory capital amounts to $$c_{P/L}L$$ and the capital constraint, $$c_{P/L}L \leq K$$, is met only if $$P \leq \frac{K - c^*L}{p} + \alpha RL$$ or:

$$P \leq \frac{c^* (L^* - L)}{p} + \alpha RL$$

(9)
It can be shown that the bank would always monitor.\textsuperscript{14} As a result, its expected profit is

$$\Pi = [pR + (1 - p)\alpha R - (1 + F)] L$$

which is increasing in $L$. Thus, the bank sets $L$ as large as possible, i.e. $L = L^*$. This however requires $P \leq \alpha RL^* \equiv \overline{P}$.

Thus we have

**Proposition 4** A CRT mechanism backed by excessive credit enhancement – condition (9) is violated for a given $L$ – undermines the bank’s monitoring incentives and entails an ex-post wealth transfer from depositors to the bank.

A capital requirement on loans conditioned on the extent of retained risk

$$c_{P/L} = c^* + p \left( \frac{P}{L} - \alpha R \right)$$

prevents such a wealth transfer, restoring efficiency.

Prudential regulation thus solves the commitment/time-inconsistency problem: Properly devised risk-based capital requirements ensure that, for any given $L$, the extent of risk retained by the bank is such that monitoring

\textsuperscript{14}The bank’s monitoring incentive-compatibility constraint is

$$PW_1 + (1 - p)W_\alpha - FL \geq PW_1 + (1 - p)W_\alpha,$$

that is,

$$W_\alpha - \frac{F}{1 - p}L \geq W_\alpha.$$

If this constraint is satisfied, securities are priced at $P_0 \equiv pLR + (1 - p)\max(\alpha LR, P)$. Clearly, if $W_\alpha > 0$, i.e. the bank is solvent in $s = \alpha$, then the monitoring incentive constraint certainly holds, because depositors do not make losses, investors break even, and hence the value loss resulting from not monitoring bears on the bank. If $W_\alpha = 0$ the constraint reduces to $W_\alpha \geq \frac{F}{1 - p}L$. Since

$$W_\alpha = \max \{ (P_0 - P) + \min(\alpha LR, P) - (L - K), 0 \}$$

$$\equiv \max \{ [p + (1 - p)\alpha] LR - p \max [P - \alpha LR, 0] - (L - K), 0 \},$$

the constraint writes as:

$$\left\{ \left[ 1 - [p + (1 - p)\alpha] R + \frac{F}{1 - p} \right] + p \max \left[ \frac{P}{L} - \alpha R, 0 \right] \right\} L \leq K,$$

which reduces to $\frac{F}{L} \geq c^* + p \max \left[ \frac{P}{L} - \alpha R, 0 \right]$, or equivalently condition (9), whenever $\frac{F}{L} \geq \alpha R$. 18
incentives are preserved. Depositors are then willing to provide the funds the bank needs for financing the efficient portfolio $L^*$, and once the funds have been raised the bank finds it profitable to engage in optimal CRT. Thus, in equilibrium, $L = L^*$ and $P = P^* \in [\underline{P}, \overline{P}]$. This also means that the regulatory capital constraint is slack.

7 Related Literature

This work is related to various strands of literature. In Duffee and Zhou (2001), the incentive for risk transfer comes from deadweight costs associated with bank insolvency. The article examines individual credit risk transfers, such as credit default swaps (CDS) and individual loan sales, and shows that as long as the asymmetric information about loan quality varies during the life of the loan, credit derivatives (CDS) dominate loan sales in circumventing lemon problems caused by banks' superior information on loan quality. However, the introduction of credit derivative markets can cause the breakdown of the loan-sale market. This is detrimental if the asymmetric information problem is one of adverse selection, whereas it may be beneficial if the problem is moral hazard, because preventing (individual) loan sales amounts to averting a decrease in monitoring. In contrast, we have shown that in itself CRT enhances monitoring. Carlstrom and Samolyk (1995) find the rationale for CRT in ceilings on banks' portfolio risks. In Morrison (2005), credit derivatives facilitate risk sharing by a risk-averse bank.

In Allen and Carletti (2006) risk aversion, and specialization across banking and insurance intermediaries, provide scope for risk-sharing between the banking and insurance sectors. The article focuses on liquidity shocks and their interaction with CRT in creating contagion between the two sectors. Wagner and Marsh (2004) hold that part of banks' credit risks may be transferred to non-banks because of the difference between banks' and non-banks' bankruptcy costs. Nicolò and Pelizzon (2007) show how various CRT instruments can signal the quality of bank loans under binding (exogenous) capital requirements. DeMarzo (2005), and earlier Boot and Thakor (1993), have shown how asset pooling and tranching may reduce informational asymmetries. Several authors have analyzed loan sales as an alternative to tradi-
tional on-balance-sheet funding that can be cheaper either because it signals asset quality (Greenbaum and Thakor, 1987) or because of exogenous reserve and capital requirements (Pennacchi, 1988; Gorton and Pennacchi, 1995).

Here we suggest a novel explanation for CRT that does not rely on bankruptcy costs, risk aversion, or regulatory constraints. In our model, optimal CRT reduces the amount of capital that the bank must stake in order for it to be incentive-compatible to monitor/screen its loans. Thus, the volume of monitored finance increases, and welfare improves. This paper therefore complements Arping (2004), who also shows that CRT can create value on incentive-based grounds, albeit for completely different reasons. In his paper, properly devised credit derivatives allow the bank to commit credibly to terminating poorly performing projects, by insulating banks from borrowers’ interim project losses. This corrects borrowers’ distorted incentives for effort and thereby increases investment project returns.

In Pennacchi (1988) and Gorton and Pennacchi (1995) regulation (capital and reserve requirements) can give banks an incentive for CRT. They study optimal individual-loan sales, i.e. the optimal portion of individual loans to be sold. This trades the benefits in terms of regulatory-cost savings, which increase with the portion sold, off against the costs of under-monitoring, which are higher, the larger the portion sold. In equilibrium, the bank sells part of the loan and chooses a suboptimal level of monitoring. The same trade-off occurs in loan syndication. Supporting empirical evidence is provided by Dennis and Mullineaux (2000). Negative repercussions of CRT on monitoring incentives are also found by Duffee and Zhou (2001), Morrison (2005), Behr and Lee (2005), Parlour and Plantin (2007), and Parlour and Winton (2007). In all these works debt financing maximizes the reward for monitoring – the monitoring-revealing outcomes are the "high" outcomes. Additional arrangements, such as CRT, can only reduce monitoring incentives.

Froot and Stein (1993, 1998) argue that otherwise risk-neutral institutions may well be averse to cash-flow volatility and engage in risk management, the argument being that agency problems constrain outside finance, making investment sensitive to the availability of internal funds. We have shown that optimal risk management in itself mitigates agency problems. The availability of outside finance expands. Finally, it is well known that collateral attenuates moral hazard and boosts the ability to raise outside
finance (Tirole, 2006). We show that optimal risk management maximizes the volume of outside finance that can be raised for any given amount of collateral – in our framework, bank capital.

Smith and Stulz (1985) show that a value-maximizing firm that raises debt will hold a hedge portfolio that shifts resources from the non-default states to the default states so as to minimize expected bankruptcy costs. However, they point out that a firm that acts on behalf of its shareholders has an ex-post incentive not to hedge, transferring wealth from debt holders to shareholders. It will then be difficult for the firm to credibly precommit to hedge, making this strategy time inconsistent. Both in Smith and Stulz and in our model, albeit for completely different reasons, risk transfer addresses the shortcomings of debt financing, and this makes it time inconsistent. We have shown that bank prudential regulation solves the time-inconsistency/commitment problem of banks’ hedging (optimal CRT).

8 Concluding Remarks

Dynamite is a very useful tool, if used properly. The same can be said of credit risk transfer. We have shown that optimal CRT creates value on incentive-based grounds: it enhances loan monitoring and expands lending. Optimal CRT instruments are based on loan portfolios rather than individual loans and have credit-enhancement guarantees. But the extent of credit enhancement needs to be precisely delimited; it must be within a defined interval. Outside that interval, monitoring incentives are undermined.

Thus, our analysis suggests that the common wisdom, supported by the previous literature, that CRT harms monitoring incentives is misleading. The problem is the bank’s incentive to credit enhance its CRT instruments to the extent required to implement the optimal contract. While insufficient credit enhancement is never profitable, we have shown that excessive credit enhancement is profitable: the bank, after collecting deposits and originating loans, has an incentive to retain too much risk, undermining its monitoring incentives. Properly designed risk-based capital requirements can restore

\[15\] Also in the banking model of Duffee and Zhou (2001), the rationale for risk transfer is costly bankruptcy. Time-inconsistency problems should then arise. However, Duffee and Zhou do not model the bank’s capital structure and assume that the deadweight costs associated with bank insolvency fall on the bank.
efficiency, making it profitable for the bank to engage in optimal CRT. But with deposit insurance and in the absence of proper regulation or other disciplining mechanisms, there is a threshold level of credit enhancement above which the quality of the assets that back CRT products deteriorates.\footnote{This may help explain some of the causes of current credit-market turmoil and the role played by the credit enhancement provided by the banks to the SIVs (Structured Investment Vehicles).}
References


Table 1

Project Return Distribution

<table>
<thead>
<tr>
<th>θ</th>
<th>Monitoring</th>
<th>Not monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr((X = R)) = 1</td>
<td>Pr((X = R)) = (\alpha)</td>
<td>Pr((X = R)) = 1</td>
</tr>
</tbody>
</table>