Credit Risk Transfer

Christine A. Parlour† Guillaume Plantin ‡

September 27, 2005

Abstract

We study the effect of endogenous financial innovation on relationship banking. Firms raise money from banks and the bond market. Banks may transfer credit risk in a secondary market to recycle their capital or to trade on private information. Liquidity in the credit risk transfer market depends on the relative likelihood of each motive for trade and affects the firm’s optimal financial structure. The conditions under which credit risk transfer markets are liquid differ from those under which they are socially optimal. The model provides testable predictions on trading volume and spreads in credit derivative markets, and prices and quantities in loan and bond markets.

† We have benefitted from comments by Doug Diamond, Denis Gromb, and seminar participants at Baruch, IHS (Vienna), the Federal Reserve Bank of New York, CFS Credit Risk Conference (Frankfurt), and the 2005 ESSFM (Gerzensee Studienzentrum). The current version of this paper is maintained at http://www.contrib.andrew.cmu.edu/˜gplantin/.

‡ Tepper School, Carnegie Mellon University and Haas School, UC Berkeley Tel: (510) 643-9391, E-mail: parlourc@andrew.cmu.edu or parlour@haas.berkeley.edu.

† Tepper School, Carnegie Mellon University Tel: (412)-268-9823, E-mail: gplantin@andrew.cmu.edu.
Introduction

In 2004, the notional value of outstanding credit derivatives was estimated at more than $4.8 trillion. In the early 1990’s, few credit derivatives were traded. Securitization consisted mostly of the transfer of senior claims on large loan portfolios. Subsequently, with the advent of Collateralized Debt Obligations (CDOs) and Credit Default Swaps (CDS), markets emerged to trade more junior exposures on less diversified risks, including first loss positions on single names. Why did the credit risk transfer (CRT) market develop in the early 1990’s, and what effect has this had on commercial banking? We characterize when a liquid CRT market arises, when a liquid CRT market is socially desirable, and provide testable predictions on the effect of such a market on prices and quantities in bond and loan markets.

In our stylized model, a firm has a positive NPV long term project that can be financed with the firm’s own equity or from two outside sources: a bank and the bond market. The optimal financing mix maximizes the firm’s return on equity subject to the resolution of a moral hazard problem: the firm may shirk which reduces the project’s success probability. A bank adds value because it can reduce a firm’s incentive to shirk. After monitoring, a bank may wish to redeploy capital and therefore sell its loans (trade in CRT). However, a liquid CRT market may also reduce a bank’s incentive to monitor as it could also sell non–performing loans. The impact of the credit risk transfer market on banks and capital markets depends on this tradeoff between a lower monitoring incentive and the increased flexibility offered by a liquid market.

In our model, banks differ from the bond market in three ways. First, banks are better able to monitor the firm than the bond market. This could be because bondholders are typically dispersed and thus individually have a smaller incentive to monitor. Second, while monitoring, the bank learns about the success or failure of the project. The assumption that lending and monitoring generates proprietary information about the borrower is common in the banking literature (Rajan, 1992), and well-documented empirically (e.g., Lummer and McConnell, 1989). In a survey on relationship banking, Boot (2000) observes, “In originating and pricing loans, banks develop proprietary information. Subsequent monitoring of the borrower yields additional private information. The existence of proprietary information may inhibit the marketability of these loans.” Finally, bank capital is scarce. Thus, the bank values liquidity and may wish to sell a loan to take advantage of private outside opportunities.

Although the rise of a liquid CRT market is socially efficient ex post, it may not be socially desirable ex ante. Because the bank prefers liquid assets, in the firm’s optimal contract, the bank’s stake is the minimum one that preserves an incentive to monitor. In particular, once monitoring has taken place, bank capital is no longer “special,” and thus the firm is indifferent as to the source of the financing. At this point it is socially efficient for the bank to sell its loan or transfer credit risk and recycle its capital. We interpret resale in the secondary market as the purchase of protection in the credit derivatives market.

Our measure of illiquidity in the credit derivatives market is the degree of adverse selection. A bank that has monitored a firm knows if the project will succeed or fail. Thus, investors do not know if the bank is selling the claim because of a private preference shock.

or because of inside information.\textsuperscript{2} If the adverse selection problem is too severe, trading in the credit derivatives market breaks down as in Akerlof (1960) and all loans are illiquid. However, if there is a high probability that the bank is selling its claim for private value motives then there is pooling and the market is liquid.

A liquid CRT market has two opposite effects on the cost of bank finance. First, interest rates on loans are smaller because the bank no longer charges a liquidity premium. However, the incentive compatible stake of the bank in the project increases, because the ability to benefit from inside trading reduces the bank’s incentives to monitor \textit{ex ante}. In sum, firms must borrow more from banks albeit at a lower price if there is an active CRT market. This is efficient if and only if the resultant cost of bank finance is smaller.

Our model has a number of testable implications. We first observe that the price of protection in a secondary market for credit risk is not the unconditional probability that the reference entity will default, rather it is the probability of default conditional on a bank’s willingness to sell. The unconditional and conditional probabilities differ systematically with a firm’s default probability, thereby generating an observed liquidity premium in the CRT market.

The rest of our testable implications relate prices and quantities in bond and loan markets after a CRT market has arisen. First, we predict that the derivatives trading volume should be larger for high grade names. Second, a larger derivatives trading volume implies that bank loans crowd out bonds in firms’ total debt. Third, a liquid derivatives market is efficient, namely increases firms’ total debt capacity, only if the firm is not too highly rated. Fourth, the pricing of bank loans is affected, and becomes more sensitive to credit risk for all ratings, even if credit derivatives are only traded for highly rated names.

Two views on the rise of credit derivatives have been proposed. First, if banks can shed credit risk, then asset-liability management is easier. Banks can quickly redeploy capital to more profitable business opportunities, and are more resilient to negative shocks. (See, for example, Greenspan, 2004, Schuermann, 2004). Following this view, liquid credit risk markets are socially beneficial. Second, banks can now unbundle balance sheet management from loan or borrower relationship management. This may reduce banks’ incentives to monitor and foster a relationship with the borrower, and ultimately create a shift from relationship banking towards transaction-oriented banking (see, for example, Effenberger, 2003; Kiff and Morrow, 2000 and Rule 2001). If relationship banking adds value, then credit risk transfer may be harmful. We demonstrate that with endogenous financial innovation, either of these cases can obtain.

Arping (2004), Duffee and Zhou (2001), and Morrison (2005) investigate the effect of credit derivatives on relationship banking. We differ from this literature in two ways. First, in our framework firms optimally mix equity, bonds and loans. We look at all sources of financing because in practice, reference entities are typically large corporations with access to diversified sources of financing.

Second, and more crucially the existing literature models the introduction of credit derivatives as an exogenous expansion of the set of feasible contracts. For instance, Arping (2004) formalizes credit derivatives as a relaxation of the limited-liability constraint. By receiving negative cash flows in case of a credit event, protection sellers reinforce the bank’s

\textsuperscript{2}Possible inside trading in the CRT market is a major concern among practitioners as reported in the popular press e.g., \textit{The Economist} Jan 16, 2003.
incentive for efficient liquidation. Morrison (2005) models credit derivatives as the introduction of noncontractible trades between the bank and protection sellers before monitoring by the bank takes place, which may be inefficient. We complement this approach by focusing instead on the real shocks that are at the origin of these institutional changes. In our framework, liquidity in the CRT market arises endogenously in response to a change in banks’ opportunity cost of capital. This is consistent with the view that financial innovations arise in response to economic shocks to financial institutions (see, e.g., Silber, 1975, 1983). We thus offer predictions on the relationship between real variables, and if and when socially desirable financial innovations arise. For example, liquid CRT markets that are not socially efficient may be viable, while some innovations that are desirable do not take place.

Our basic model of moral hazard is adapted from Holmstrom and Tirole (1997). The notion that banks value liquidity and acquire proprietary information through monitoring follows Breton (2003). He introduces a similar tension between monitoring and liquidity, in a different context. In non-banking contexts, Aghion, Bolton, and Tirole (2004), and Faure-Grimaud and Gromb (2003) develop models in which an impatient agent markets claims to her future output. The informational efficiency of the market is an important incentive device for the agent. By contrast in our model, liquidity can make incentives more difficult to enforce because in a liquid market, the bank can trade with uninformed agents.

1 Model

In an economy with three dates, \( t = 0, 1, \) and \( 2 \), there are three classes of agents: a firm, a bank, and a large pool of competitive investors. All agents are risk neutral and are protected by limited liability.\(^3\) Neither the risk neutral investors nor the firm discount future cash flows.

A firm owns the rights to a two period constant returns to scale project. Every dollar invested in the project at \( t = 0 \) has a random date 2 payoff of \( R \) with probability \( p \) or 0 with probability \( 1 - p \). The firm also has initial equity or net assets, \( A \). In addition to investing its own assets in the project, the firm can solicit additional funds from the bank and the outside investors. To do so, the firm offers each group a renegotiation proof contract. The firm’s objective is to maximize expected return on equity.

Some actions taken by the firm are not observable to outside investors. In particular, after raising outside funds, the firm may “shirk” and derive a private benefit \( B_F \) per dollar invested at \( t = 1 \). In this case, the project fails and pays off 0 at date 2. This moral hazard problem means that financial structure matters.

Active monitoring at \( t = 0 \) by the bank reduces the firm’s private benefit per unit of investment from shirking to \( b_F < B_F \). Monitoring is costly for the bank as it loses a private benefit \( B_B \) per dollar of the project associated with not monitoring.\(^4\) A benefit of monitoring is that, through its relationship with the firm, the bank acquires private information about the project. For simplicity, we assume that private information is perfect: if the bank monitors, it learns the project’s outcome at date 1. To capture the idea that bank capital

\(^3\)Specifically, we assume that all contractible payments are bounded below by a finite amount that we normalize to 0.

\(^4\)Alternatively, one could model the bank as incurring a private monitoring cost. The current formulation is the simplest one that ensures all constraints bind at the optimum.
is costly, we assume that the bank discounts future cash flows. More precisely, the bank values cash flows at

\[ \pi_B(c_0, c_1, c_2) = E(c_0 + \delta_1 c_1 + \delta_1 \tilde{\delta}_2 c_2), \]

where \( \delta_1 \in (0, 1) \), and \( \tilde{\delta}_2 \) is a two point random variable whose realization at date 1 is

\[ \tilde{\delta}_2 = \begin{cases} \delta & \text{with prob. } q \\ 1 & \text{with prob. } 1 - q. \end{cases} \]

The bank’s realization of \( \tilde{\delta}_2 \) at date 1 is private information.\(^5\) This stochastic discount factor proxies for unanticipated changes in the opportunity cost of carrying outstanding loans. For example, a financial institution could receive private opportunities to invest with other borrowers or even non lending business.\(^6\)

After the bank has received her discount factor shock and learned the realization of the project, she can sell her stake in the project in a CRT market (if it is liquid). The counterparties are the risk neutral investors.

The time line is presented in figure 1.

<table>
<thead>
<tr>
<th>Contracts written and Firm Invests</th>
<th>Bank learns: (i) opp. cost (( \tilde{\delta}_2 ) is ( \delta ) or 1) (ii) if project succeeds</th>
<th>Possible CRT trading</th>
<th>Claims Pay off</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bank chooses monitor or not</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Firm shirks or not</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t = 0 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t = 1 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t = 2 )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Sequence of events

We have assumed that the bank may monitor before CRT trading. This assumption is for simplicity: An additional moral hazard problem at period 2, or possible CRT trade at date 0, would complicate the analysis without offering any further insights in our environment of complete contracts. Morrison (2005) studies a model with CRT followed by monitoring in an incomplete contracting environment. He demonstrates that complete contracts would remove all inefficiencies.

We assume that all random variables are independent. To ensure interior solutions we impose additional parameter restrictions when we solve for the optimal contracts.

## 2 Characterization of the Optimal Contracts

Before we characterize the optimal contracts a firm offers to the bank and outside investors, we determine the price at which CRT can occur: contract specifics depend on this.

---

\(^5\)We have assumed that the realization of bank’s discount factor at \( t = 2 \) is binary for simplicity. The qualitative results go through if the bank’s outside opportunities are drawn from a known continuous distribution.

\(^6\)We elaborate further on opportunity costs of carrying loans in Section 3.
2.1 Prices in the CRT Market

At $t = 1$, the bank may sell the loan in the CRT market. At this time it has two pieces of private information. First, the bank’s opportunity cost of investment ($\tilde{\delta}_2$) has been realized. Second, it knows if the firm’s project has succeeded or failed. A failed project pays off 0. Thus, a bank with a failed project will sell it at any positive price. However, if the only motive for trade in the CRT market is to dispose of failed projects, then the price must be zero. We term such a market illiquid. Alternatively, the price could be such that a bank with a liquidity shock would also be willing to sell the loan. We deem such a market to be liquid.

**Definition 1** A CRT market is liquid if a bank sells both failed and successful claims. A CRT market is illiquid if a bank only sells failed claims.

If the market is liquid, then investors believe that the bank is selling either because the project failed (which occurs with probability $1 - p$) or that the project succeeded but the bank received an attractive outside opportunity. This occurs with probability $pq$. Thus, if the market is liquid, outside investors value one promised date 2 dollar at a price $r$ where

$$r = \frac{pq}{1 - p + pq}.$$  

Notice that $r < p$, the unconditional probability that the project succeeds. The price, therefore, incorporates an adverse selection discount.

The price at which outside investors are willing to buy future claims can be used to determine the price of protection. In the CDS market, conditional on a credit event, the protection buyer receives either the par value of the claim and delivers the securities or the difference between the notional amount and the post default market value of the reference asset. In this model, if the firm defaults the claims are worthless. Thus, for both cash or physical settlement the default payment is zero. Consider an obligation to pay $1 at $t = 2$ if the firm does not default and zero otherwise. If a CRT market exists, this obligation is worth $r$. Thus, a claim that pays $1 if the firm does default and 0 otherwise, must be worth $1 - r$. We therefore interpret $1 - r = \frac{(1-p)}{1 - p + pq}$ as the price of protection.

It is immediate that in this framework, the price of protection for a fixed investment size is not the default probability $(1 - p)$. Rather, it is the probability that the firm defaults conditional on the bank’s willingness to sell the claim. This is always weakly higher than the default probability because of the adverse selection discount.

Figure 2 depicts the default probability and the price of protection.
Figure 2: The price of protection and default probabilities

Clearly, for $p = 1$ or $p = 0$, there is no informational asymmetry and the price of protection is equal to the default probability. However, for $p \in (0, 1)$ the price of protection is strictly greater than the default probability. Let $\Delta(p)$ denote the difference between the price of protection and the default probability.

\[
\Delta(p) = (1 - r) - (1 - p) = \frac{(1 - p) p (1 - q)}{1 - p + pq} \geq 0
\]

Thus, $\Delta(p)$ is the liquidity premium, or that part of the price of protection that is not explained by default probabilities.

**Lemma 1** If a CRT market is liquid, then
(i) The liquidity premium, $\Delta(p)$ is concave in $p$, the probability of no-default.
(ii) The ratio of the liquidity premium to the default probability, $\frac{\Delta(p)}{1 - p}$ is increasing, and convex in the probability of no-default.

Thus, the liquidity premium is increasing in the credit rating $p$ for a low $p$, decreasing in the credit rating for a high $p$. However, the liquidity premium as a fraction of the default probability is always increasing in the credit rating.

### 2.2 Characterization of Optimal Contracts

We characterize renegotiation proof contracts offered by the firm at $t = 0$ to the different creditors. Such contracts differ if participants believe the CRT market will be liquid. We also note that any renegotiation proof contract does not include any clause restricting date 1 trade in the CRT market. This is because the bank is hit by a private shock and receives private information after monitoring has taken place. How the bank and investors
trade in the secondary market is therefore irrelevant for the firm at date 1. So, any clause deviating from \textit{ex post} efficient CRT activity would not be renegotiation proof. We note that empirically, such clauses do not appear in loan contracts. This distinguishes our paper from both Morrison (2005) and Aghion, Bolton, and Tirole (2004). The former assumes that agents cannot contract on CRT, and the latter assume full commitment rather than renegotiation proof contracts.

Recall the firm has assets $A$, and owns a project which generates a return per unit of investment of $R$ with probability $p$ and 0 with probability $1 - p$. For a given investment size, $I$, the firm pledges a portion of the final payoff to the bank, $R_B I$ to secure a loan of size $L$, and another portion $R_M I$ to the market to secure a bond issue of size $M$. The firm retains a portion $R_F I$ and maximizes expected profits $\pi^F = p R_F I - A$ by choosing $\{I, L, M, R_F, R_B\}$.

First, suppose that bank monitoring is inefficient, or that $B_B$ is sufficiently high and the firm only borrows from the bond market. Then, the firm maximizes surplus subject to the incentive-compatibility constraint (Equation 1) and investors’ participation constraint (Equation 2). These are

\begin{align}
  p R_F I &\geq B_F I, \\
p(R - R_F) I &\geq I - A.
\end{align}

If these constraints bind at the optimum then the IC (Equation 1) determines $R_F$, while the market’s participation constraint (Equation 2) determines the investment level $I(p)$.

\textbf{Lemma 2} If the firm borrows only from outside investors, then if $\max[B_F, 1] < p R < 1 + B_F$ the expected profit of the firm, $\pi^F(p)$, the optimal investment $I(p)$ and the optimal outside investment, $M(p)$ are

\begin{align*}
  \pi^F &= \left(\frac{p R - 1}{1 - p R + B_F}\right) A \\
  I(p) &= \frac{A}{1 - p R + B_F} \\
  M(p) &= \left(\frac{p R - B_F}{1 - p R + B_F}\right) A.
\end{align*}

Assume now that bank monitoring is valuable. To characterize the optimal contracts when a bank monitors, we consider two cases depending on whether an active CRT market can exist. If there is no active CRT market, the firm chooses a contract:

\begin{align*}
  \max_{I, L, M, R_F, R_B} &\{p R_F I - A\} \\
  \text{s.t.} &
  p R_F &\geq B_F \\
  \delta_1 E \delta_2 p R_B &\geq B_B \\
  \delta_1 E \delta_2 p R_B I &\geq L \\
  \frac{p (R - R_F - R_B)}{R_M} I &\geq M
\end{align*}
Conditions (3) and (4) are the incentive compatibility constraints of the firm and the bank, respectively. Equation (3) ensures that the firm, if monitored, will not shirk. If the firm does not shirk, then it receives a payoff of \( pR_F I \), whereas if it shirks and is monitored by the bank, the payoff is \( b_F I \). Condition (4) ensures that the bank’s payoff is higher if it monitors the firm than if it does not. Specifically, if the bank does not monitor, it consumes private benefits proportional to the size of the project, or \( B_B I \). If the bank does monitor the firm, then it receives a promised payoff of \( R_B I \) with probability \( p \). The time \( t = 1 \) value of this expected payout depends on the realization of the bank’s discount factor between \( t = 1 \) and \( t = 2 \). The date 0 expected value of the discount factor is \( E\delta_2 \). The bank discounts cash flows between \( t = 1 \) and \( t = 0 \) at \( \delta_1 \).

**Lemma 3** If there is no CRT market and if the firm borrows from both the bank and outside investors, then if 
\[
1 + B_B \frac{1-\delta_1 E\delta_2}{\delta_1 E\delta_2} < pR < 1 + B_B \frac{1-\delta_1 E\delta_2}{\delta_1 E\delta_2} + b_F,
\]
the profits to the firm, \( \pi^F \), the project size \( I(p) \), the amount borrowed from the bank, \( L(p) \), and the amount borrowed from the market, \( M(p) \), are

\[
\pi^F = \left( \frac{b_F}{1 - pR + b_F + B_B \frac{1-\delta_1 E\delta_2}{\delta_1 E\delta_2}} - 1 \right) A
\]

\[
I(p) = \frac{A}{1 - pR + b_F + B_B \frac{1-\delta_1 E\delta_2}{\delta_1 E\delta_2}}
\]

\[
M(p) = I(p)(1 - B_B) - A
\]

\[
L(p) = B_B I(p)
\]

Now, suppose that a CRT market exists, and a bank can sell claims to \( t = 2 \) cash flows at a price \( r \) at date 1. The firm then solves:

\[
Max_{I,L,M,R_F,R_B} \quad \{ pR_F I - A \}
\]

s.t.

\[
pR_F \geq b_F
\]

\[
\delta_1 pR_B \geq B_B + \delta_1 r R_B
\]

\[
L \leq \delta_1 pR_B I
\]

\[
p(R - R_F - R_B) I \geq M
\]

\[
I \leq M + A + L
\]

\[
I, L, M, R_F, R_B \geq 0
\]
failed or if it receives a discount factor shock. With probability \((1 - p + pq)\) the bank can sell its claim at price \(r = \frac{pq}{1 - p + pq}\). With probability \(p(1 - q)\), it will not sell its claim but value it at a discount rate of \(q\). Thus, the expected \(t = 1\) value of a cash flow promised at \(t = 0\) is \(1\) which the bank discounts to \(t = 0\) at \(\delta_1\). If the bank shirks (the righthand side of (9)), it receives private benefits of \(B_B I\). In addition, it knows that the project has failed. It can therefore sell its promised payment of \(R_B I\) at a price of \(r\) in the CRT market. The \(t = 0\) value of this sale is \(\delta_1 r R_B I\).

The change in expected discount rate affects a bank’s behavior through both the participation and the incentive compatibility constraint. It is instructive to compare these new constraints to those that obtain without a CRT market, namely constraints (4), and (5). The bank’s participation constraint is easier to meet with this higher discount factor because the cost of bank capital is lower: The price of the loan no longer features the liquidity premium \(E \delta_2\). However, the impact of a liquid market on the incentive compatibility constraint is ambiguous. First, the bank does not require a liquidity premium on date 2 cash flows. This reduces the cost of incentives. However a bank who does not monitor and consumes private perquisites can now sell the worthless loan in the CRT market at the pooling price. This makes shirking more attractive and thus the IC may be more difficult to satisfy. We demonstrate in Section 3 that the optimal solution to this problem is to take out proportionally larger loans from the bank. In sum, with an active CRT market, the firm needs to borrow larger quantities from the bank, but each dollar is borrowed at a lower cost.

**Lemma 4** If there is a liquid CRT and if 
$$\max \left( b_F + B_B \left( \frac{1}{\delta_1} \right) \frac{p}{p - r} ; 1 + B_B \frac{p}{p - r} \frac{1 - \delta_1}{\delta_1} \right) < pR < 1 + B_B \frac{p}{p - r} \frac{1 - \delta_1}{\delta_1} + b_F,$$
then the surplus to the firm, \(\pi_F\), the size of the project, \(I\), \(M\), \(L\) are

\[
\pi_F = \left( \frac{b_F}{1 + b_F - pR + B_B \left( \frac{p}{p - r} \right) \left( \frac{1 - \delta_1}{\delta_1} \right)} - 1 \right) A
\]

\[
I(p) = A \left( \frac{1}{1 + b_F - pR + B_B \left( \frac{p}{p - r} \right) \left( \frac{1 - \delta_1}{\delta_1} \right)} \right)
\]

\[
L(p) = B_B \left( \frac{p}{p - r} \right) I(p)
\]

\[
M(p) = I(p) \left( \frac{(1 - B_B)p - r}{p - r} \right) - A
\]

Lemma 4 and 3 presents solutions to the model under the respective parameter restrictions:

$$1 + B_B \frac{1 - \delta_1 E \delta_2}{\delta_1 E \delta_2} < pR < 1 + B_B \frac{1 - \delta_1 E \delta_2}{\delta_1 E \delta_2} + b_F,$$

if there is no CRT, and

$$\max \left( b_F + B_B \left( \frac{1}{\delta_1} \right) \frac{p}{p - r} ; 1 + B_B \frac{p}{p - r} \frac{1 - \delta_1}{\delta_1} \right) < pR < 1 + B_B \left( \frac{p}{p - r} \right) \frac{1 - \delta_1}{\delta_1} + b_F.$$
if there is a CRT. In both cases, the left-hand inequality ensures that the project NPV is high enough so that the firm can borrow from the bank and the market and that bank monitoring is feasible. The right-hand inequality bounds the NPV so that the optimal investment size is finite. When we perform comparative statics we assume that these conditions hold; effectively \( B_B \) sufficiently small and \( b_F \) sufficiently large. For a fixed \( R \), the conditions may restrict default probabilities to a small range. Empirically, Hamilton (2001) estimates average five year cumulative default rates for investment grade bonds as 0.82% and 18.56% for speculative grade bonds. Thus, the interval of default probabilities \( 1 - p \) that we consider is small. It is therefore reasonable to assume that the parameter restrictions hold over the whole interval.

3 Efficiency and CRT

By assumption, the bank and bondholders are always held to their participation constraints. Thus, social surplus is maximized when the firm makes the highest profit. Given the constant returns to scale technology, profits are maximal when the size of the project, \( I(p) \) is maximal. Equivalently, social efficiency demands that a firm borrows as much as possible. Recall,

\[
I(p) = A \left( \frac{1}{1-pR+bF+B_B \frac{1}{\delta_1 \delta_2}} \right) \quad \text{if there is no CRT}
\]

\[
\frac{1}{1-pR+bF+B_B \frac{1}{p-r} \frac{1}{\delta_1}} \quad \text{if there is a CRT}.
\]

The existence of CRT has two countervailing effects on investment size. On one hand, with an active CRT, the bank’s participation constraint is easier to satisfy. This is because the bank no longer demands a liquidity premium as it discounts \( t=2 \) expected cash flows at \( \delta_1 \) not \( \delta_1 E \delta_2 \). Thus, the firm can compensate the bank at \( t=0 \) with a lower expected return. Let \( r_B(p) = \frac{pR}{L}I(p) \) denote the time 0 expected return on a bank loan.

**Lemma 5** The expected rate of return on bank loans, \( r_B(p) \) is lower if the CRT market is liquid.

On the other hand, if there is a liquid CRT, it is more difficult to make monitoring incentive compatible. This is because a bank can sell non performing loans at the pooling price. To mitigate this effect, the bank’s stake in the time 2 payoff if there is a CRT has to be higher than if there is not. Bank capital is more expensive than bond financing, and the former crowds out the latter, as the surplus pledgeable to bondholders is reduced by this increase in expensive capital. Thus, the proportion of public debt is lower. Let \( f_B(p) = \frac{B_B I(p)}{L} \) denote the fraction of the project funded by bank debt. Let \( f_M(p) = \frac{M(p)}{I(p)} \) denote the fraction of the project funded by market debt.

**Lemma 6** (i) The ratio of bank debt over total project size, \( f_B(p) \), is higher if the CRT market is liquid. (ii) The ratio of public debt over total project size, \( f_M(p) \), is smaller if the CRT market is liquid.
While the per unit cost of bank capital is lower, proportionally more has to be solicited to ensure that the bank monitors. If the net effect is a reduction in the total surplus paid to the bank, then investment, \( I(p) \), increases. Conversely, if there is an increase in the total surplus accruing to the bank, then \( I(p) \) decreases.

**Proposition 1** A liquid CRT market is efficient if and only if \( r_B(p)f_B(p) \), the expected cost of a bank loan per unit investment, is smaller if the market is liquid.

To see this, it is useful to observe that optimal investment can be expressed as

\[
I = \frac{A}{1 - pR + bF + r_B(p)f_B(p)}.
\]

Clearly, investment size is decreasing in both \( bF \) and \( r_B(p)f_B(p) \). These can be interpreted as informational rents accruing to the firm and the bank. While CRT does not affect the firm’s informational rent, it does affect the bank’s. CRT is socially desirable only if it reduces the bank’s rent per unit investment. As we have observed, ex post efficiency conflicts with ex ante efficiency and a liquid CRT market may reduce investment. Thus, the parameters under which CRT arises and the parameters under which CRT is ex ante desirable differ.

**Proposition 2**

(i) A liquid CRT market exists if and only if the possible discount factor shock (\( \delta \)) is sufficiently small so that \( \delta \leq \frac{pq(1-p)}{1-p+pq} \).

(ii) A liquid CRT market is socially efficient if and only if the possible discount factor shock (\( \delta \)) is sufficiently small so that \( \delta \leq \frac{(1-q)(\delta_1-p)}{(1-p)^2-q(\delta_1-p)} \).

The first condition of Proposition 2 states that for the CRT market to be liquid, the pooling price \( r = \frac{pq}{1-p+pq} \) must be high enough so that banks who receive discount factor shocks are willing to sell at this price. Or, the beliefs of market participants are on average correct.

Interestingly, Acharya and Johnson (2005) find that CDS markets for corporate entities that have a large number of banking relationships tend to be more liquid. The logic of our model suggests that multiple banking relationships would improve liquidity. If one bank out of many buys credit protection for an entity the probability that the bank is trading for private motives must be higher. Conversely, if all the entity’s bankers sell at the same time trade is probably occurring because of an information event. Thus, multiple lenders reduce the likelihood that investors are faced with adverse selection.

Further, notice that for a fixed value of \( \delta \), a CRT market is ceteris paribus easier to sustain for higher rated names. The fact that CRT markets are easier to sustain for high credit rated firms (high \( p \)) corresponds to casual empiricism. Active CDS markets exist for highly rated names but not for those below investment grade. For example, a 2003 study by Fitch cited in Bomfim (2005) estimates that 90% of CDS were written on investment grade bonds, with 28% BBB, 28% A, 15% AA and 21% AAA. Only 8% were below investment grade.

Alternatively, for a fixed default probability, financial innovation arises when the bank’s private cost of bearing risks until maturity is greater than the liquidity premium in the
CRT market. Thus, we should expect to see more risk transfer vehicles arise when banks are faced with a higher opportunity cost of lending (lower $\delta$). This observation is the basis for our comparative statics results in Section 4.

The second condition indicates that for higher rated names, a CRT market is more likely to be inefficient. Indeed,

**Proposition 3** There exists a $\hat{p} < 1$, so that for $p > \hat{p}$ any liquid CRT market is inefficient.

In our model, financial innovation is efficient if more is invested. As we have indicated, the price of bank loans always decreases with the advent of a CRT market, as banks no longer demand a liquidity premium. This effect is independent of the credit rating of the underlying name and only depends on the bank’s opportunity cost of capital ($\delta$). By contrast, as the credit quality of a firm increases, increasing amounts of bank loans have to be solicited with the advent of a CRT market. Firms with high credit quality are very valuable in the CRT market and thus there is a proportionally larger benefit to shirking. For large enough $p$, the price effect is outweighed by the quantity effect leading to inefficiency.

In Figure 3, we illustrate the two conditions of Proposition 2. The condition for efficiency is the threshold below which aggregate investment is larger with a CRT. There are four possibilities in this economy.

Figure 3: Efficiency and Liquidity

Figure 3 illustrates that there can be too much or too little financial innovation. Financial institutions may inefficiently lay off low risk projects which leads to lower overall investment, or may inefficiently retain more risky projects when their opportunity cost is low.
4 Cross-Sectional Implications

In our model, \( \delta \) is the bank’s opportunity cost of bearing credit risk. Thus, it represents both the outside opportunities presented to the bank and the regulatory constraints imposed on the banks. A low \( \delta \) corresponds to either profitable outside lending opportunities and tight regulatory controls, or both. In the United States, the Basel Capital rules requiring banks to maintain capital reserves of 7.25% of loans were adopted in 1989. The reserve requirement was increased to 8% in 1992. In 1991, the Federal Deposit Insurance Corporate Improvement Act (FDICIA) further tightened regulatory control over commercial banks. Finally, the Reigle-Neal Act removed barriers to interstate banking in 1994. Thus, in contrast to the late 1980’s; the early to middle 1990’s were a time when bank’s use of capital was constrained by legislation while new banking opportunities emerged. We interpret the late 1980’s as a period when \( \delta \) was high and the early to middle 1990’s as a period when \( \delta \) was low.\(^7\)

We posit that the rise of credit derivatives was triggered by this change in \( \delta \), and consider changes in market aggregates that must also have occurred. Formally, consider a population of firms who differ only in credit ratings. Assume that the success probabilities, \( p \), are distributed on \([p, \bar{p}]\) according to \( G(\cdot) \). We compare the cross-sectional variations of firms’ financial structure for two different values of \( \delta \), denoted \( \hat{\delta} \) and \( \bar{\delta} \), where

\[
0 < \hat{\delta} < \bar{\delta} < 1.
\]

Here, \( \hat{\delta} \) is the opportunity cost of lending in the mid 1990’s while \( \bar{\delta} \) is that of the late 1980’s. As the CDS market emerged in the 1990’s, we assume that no CRT market is feasible for \( \bar{\delta} \), while for \( \hat{\delta} \) there is a liquid CRT market for all \( p \geq p^* \), where \( p^* \in [\hat{p}, \bar{p}] \). This situation is depicted in Figure 4. As already noted, we correctly predict that CRT arises only for sufficiently highly rated entities.

Figure 4 here

**Proposition 4** Suppose \( \delta \) decreases from \( \bar{\delta} \) to \( \hat{\delta} \). Consider two firms that differ by their success probabilities \( p'' > p' \), where \( p'' \in [p^*, \bar{p}] \), and \( p' < p^* \). Then, the difference in the expected return on bank loans is larger for \( \hat{\delta} \) than \( \bar{\delta} \). Or,

\[
r_B(p' | \hat{\delta}) - r_B(p'' | \hat{\delta}) > r_B(p' | \bar{\delta}) - r_B(p'' | \bar{\delta}).
\]

This result follows directly from the fact that the bank’s participation constraint is binding. The proposition could, alternatively, be interpreted as a statement about bank’s internal cost of capital or the discount factor a bank applies to loans with specific default probabilities. If there is no CRT, then the difference in discount factors between a firm with \( p' \) and \( p'' \) is

\[
\delta_1(p'' - p')(q\delta + 1 - q) \quad \text{for} \quad \delta \in \{\hat{\delta}, \bar{\delta}\}.
\]

\(^7\)We focus on a decrease in \( \delta \) to simplify the analysis. An increase in \( q \) yields similar results.
If a CRT market arises, then the bank’s discount factor increases for borrowers with a rating above $p^*$ because the illiquidity discount $(q\delta + 1 - q)$ disappears. For $p \leq p^*$ the discount factor is now smaller as a lower $\delta$ commands a higher liquidity premium. Empirically, this implies that the bank loan pricing schedule is more sensitive to borrower credit quality in the late 1990s than the late 1980s. More precisely, the slope of the promised return on bank loans $(\frac{1}{\delta} - 1)$ as a function of the rating $p$ should have steepened after the rise of credit derivatives, with no change in the distribution of credit losses. Schuermann (2004) finds some evidence in support of this phenomenon. He finds that an estimate of the slope for bank loan pricing schedules has steepened during the 90s, while its equivalent for bonds pricing schedules has flattened, if anything.

The proportion of bank financing in the economy should also have become more sensitive to credit rating.

**Proposition 5** Suppose $\delta$ decreases from $\bar{\delta}$ to $\underline{\delta}$. Then, the fraction of investments that is financed with bank loans becomes more sensitive to the credit rating of firms. Or, for $p' < p^* < p''$,

$$f_B(p'', \delta) - f_B(p', \delta) < f_B(p'', \bar{\delta}) - f_B(p', \bar{\delta}).$$

In our model, for $p > p^*$, this fraction shifts from flat for $\delta$, to increasing and convex for $\bar{\delta}$.

**Proposition 6** Suppose $\delta$ decreases from $\bar{\delta}$ to $\underline{\delta}$. Then, the fraction of investments that is financed with bonds decreases for any rating. Or,

$$\forall p \in [\underline{p}, \bar{p}], f_M(p, \underline{\delta}) < f_M(p, \bar{\delta}).$$

Empirically, demand factors not captured in our model may have spurred bond finance (for example the introduction of the euro in Europe). Such unmodelled forces would lead to an increase in the proportion of bond financing over the period. Thus, the empirical translation of this proposition would rather be that the fraction of bond finance has grown less quickly in countries in which banks use credit derivatives more intensively other things being equal.

Finally, trade in the CRT market is efficient for a given name if and only if, as a result, the cost of bank finance per unit investment decreases for this name. Proposition 7 gives a sufficient condition for the aggregate inefficiency of CRT that may be easier to test because it relies solely on changes in total investment capacities across credit ratings.

**Proposition 7** Suppose that $\delta$ decreases from $\bar{\delta}$ to $\underline{\delta}$. Then, a sufficient condition for CRT to be socially inefficient is if the ratio of total investment in traded names over non–traded names increases. Or,

$$\frac{\int_{p}^{p^*} I(p, \underline{\delta}) \, dG(p)}{\int_{p}^{p^*} I(p, \bar{\delta}) \, dG(p)} < \frac{\int_{p}^{p^*} I(p, \underline{\delta}) \, dG(p)}{\int_{p}^{p^*} I(p, \bar{\delta}) \, dG(p)}.
\tag{11}$$

This result stems simply from the fact that, absent CRT, the investment capacity becomes less sensitive to the credit rating ($p$) for a low $\delta$. Thus, inequality (11) is reversed.
absent CRT. Therefore, if this inequality holds, CRT has increased aggregate investment for firms rated above $p^*$ and must be efficient. Implicit in Condition 11 is the idea that CRT arises due to an increase in the cost of bank capital. This increased cost leads to a decrease in total investment across all projects. Or, $I(p \mid \delta) > I(p \mid \tilde{\delta})$. An immediate implication of this is that (absent a CRT) aggregate defaults in the economy ($\int_\mathbb{P} [1 - p]I(p \mid \delta))$ have decreased if the cost of bank capital has increased. This observation can also extend to an economy in which CRT arises. In particular, if a CRT market is inefficient, then aggregate investment as a function of credit rating is smaller than it would be without CRT. Thus, if a change in bank capital fosters the growth of the CRT market, aggregate defaults in the economy must also have fallen if this market is inefficient.

Further, as the amount of market debt is always a (weakly) decreasing proportion of the total amount borrowed, aggregate losses on market debt are lower in the economy after the advent of a CRT even though its existence is inefficient. This is because the rise of the CRT market is due to a real shock on bank’s opportunity cost of capital.

\textbf{Proposition 8} Suppose $\delta$ decreases from $\tilde{\delta}$ to $\delta$, and CRT is inefficient, then aggregate losses on market debt are lower in the economy. Or,

$$\int_\mathbb{P} (1 - p)M(p \mid \tilde{\delta})dG > \int_\mathbb{P} (1 - p)M(p \mid \delta)dG$$

Thus, aggregate defaults on market debt decrease even though the level of investment is not socially desirable. This is not the result of a change in underlying default probabilities.

\section{Conclusion}

We have analyzed the recent rise of credit derivatives with a simple model of endogenous financial innovation. Financial institutions innovate when they find that trading a risk is preferable to bearing it. Their risks become liquid when the benefits from freeing up regulatory capital overcome the informational cost of shedding risks. In the particular case of default risk, we posit that because banks’ opportunity cost of loans has increased, markets have evolved which allow banks to separate balance sheet management and borrower relationship management.

Because a liquid CRT market implies \textit{ex ante} inefficiencies, we predict possible excessive CRT trading in high rated names, and insufficient liquidity of lower rated names. The rise of more risk-based capital requirements for financial institutions should contribute to redeploy CRT liquidity where it is the most desirable. According to our theory, a decrease in the opportunity cost of lending to high quality borrowers should reduce trading volume in the secondary market for such borrowers. Conversely, a higher cost of lending to low rated entities should increase CRT liquidity for these entities. Thus, the implementation of the “Basel II” revised capital framework for banking organizations, and the development of more risk based internal capital budgeting rules within banks should imply a socially efficient shift of trading volume towards lower rated names in credit derivatives markets.
6 Appendix

Proof of Lemma 1
Recall, \( \Delta(p) = \frac{(1-p)p(1-q)}{1-p(1-q)} \). Thus,
\[
\frac{d\Delta(p)}{dp} = \frac{(1 - 2p + p^2 (1 - q)) (1 - q)}{(1 + p (-1 + q))^2}
\]
Here, \( \Delta(p) = \frac{p(1-q)}{1-p+pq} \). Clearly
\[
\frac{d \left( \frac{\Delta(p)}{1-p} \right)}{dp} = \left( \frac{p (1 - q) (1 - q)}{(1 - p + pq)^2} \right) + \frac{1 - q}{1 - p + pq} \geq 0
\]
\[
\frac{d^2 \left( \frac{\Delta(p)}{1-p} \right)}{(dp)^2} = \frac{2 p (1 - q) (1 - q)^2}{(1 - p + pq)^3} + \frac{2 (1 - q) (1 - q)}{(1 - p + pq)^2} \geq 0.
\]

Proof of Lemma 2, 3, 4
For each of these cases, all the constraints are binding. The stated results follow.

Proof of Lemma 5
The promised return per dollar invested is \( \frac{pR}{L} \). Thus, if there is a CRT, this is \( \frac{pR}{L} = \frac{1}{\delta_1} \). If there is no CRT market then \( \frac{pR}{L} = \frac{1}{E\delta_2\delta_1} \).
The result follows from \( E\delta_2 < 1 \).

Proof of Lemma 6
(i) The ratio of bank debt over total project size if there is an active CRT Market is
\[
\frac{L^{CRT}}{I^{CRT}} = \frac{pB_B}{p - r},
\]
and if there is not,
\[
\frac{L^{no}}{I^{no}} = B_B.
\]
Thus \( \frac{L^{CRT}}{I^{CRT}} \geq \frac{L^{no}}{I^{no}} \) if \( p \geq p - r \), which always holds.

(ii) The ratio of market debt over total project size, if there is no CRT market is
\[
\frac{M^{no}}{I^{no}} = pR - b_F - B_B \frac{1}{\delta_1 E\delta_2},
\]
and if there is one,
\[
\frac{M^{CRT}}{I^{CRT}} = pR - b_F - B_B \frac{p}{\delta_1(p - r)}.
\]
\( \frac{M^{no}}{I^{no}} \geq \frac{M^{CRT}}{I^{CRT}} \) if \( E\delta_2 > \frac{p-r}{p} \), which always holds.
Proof of Proposition 2

(i) A bank will sell the loan on a failed project at any price \( r \geq 0 \). A bank with a successful project values it at \( \delta R \), thus will sell it if \( rR \geq \delta R \), or \( r \geq \delta \). If \( r < \delta \), then only the banks who know that the project is unsuccessful will sell, thus claims are worth 0. If \( r \geq \delta \), then with probability \( 1 - p \), the bank knows the project is a failure, and with probability \( pq \), the project was a success, but the bank got a shock. Thus, if \( r \geq \delta \), the expected value of $1 promised at date 2 is \( \frac{pq(1-p)}{pq(1-p)} \).

(ii) stems directly from

\[
I = A \begin{cases} 
1 - \frac{\delta_1 E_2}{\delta_1 E_2} & \text{if there is no CRT} \\
1 - \frac{\delta_1 E_2}{\delta_1 E_2} & \text{if there is a CRT}, 
\end{cases}
\]

Thus, the existence of a CRT is efficient (results in a larger investment) if and only if \( \frac{pR_B}{1-L} \) is smaller.

Proof of Proposition 3

Recall, \( r = \frac{pq}{1-p+pq} \). Thus, \( r \) is increasing and convex in \( p \) through the origin. Let \( \delta^* \) be the maximum \( \delta \) for which a CRT market is socially efficient. Thus, \( \delta^* = \frac{(1-q)(\delta_1 - p)}{(1-p-q(\delta_1 - p))} \).

Notice, when \( p = 0 \), \( \delta^* = \frac{(1-q)\delta_1}{1-q\delta_1} \). Further, when \( p = \delta_1 \), \( \delta^* = 0 \)

\[
\frac{d\delta^*}{dp} = -\frac{(1-q)(1-\delta_1)}{(1-p+q(p-\delta_1))^2} < 0. 
\]

Thus, there is a unique intersection point less than one.

Proof of Proposition 4

If there is no CRT, then \( r_B(p, \delta) = \frac{1-\delta_1 E_2}{\delta_1 E_2} \). If there is a CRT, then \( r_B(p, \delta) = \frac{1-\delta_1}{\delta_1} \).

Suppose that \( \delta = \delta^* \), then there is no CRT (by assumption). Whereas, if \( \delta = \delta^* \), then for \( p > p^* \) a CRT market is feasible. Comparison of the expected returns yields the result.

Proof of Proposition 5

From Proof of Lemma 6, the ratio of bank debt over total project size if there is an active CRT Market is

\[
\frac{L^C}{I^C} = \frac{pB_B}{p-r},
\]

and if there is not,

\[
\frac{L^c}{I^c} = B_B.
\]
Proof of Proposition 6

From Proof of Lemma 6, the ratio of market debt over total project size, if there is no CRT market is

\[ \frac{M^{no}_{I^{no}}}{I^{no}} = pR - bF - B\frac{1}{\delta_1E\delta_2}, \]

and if there is one,

\[ \frac{M^{CRT}_{I^{CRT}}}{I^{CRT}} = pR - bF - B\frac{p}{\delta_1(p - r)}. \]

\[ \blacksquare \]

Proof of Proposition 7

In the absence of a CRT, \( I(p | \bar{\delta}) > I(p | \delta) \). Further, \( \frac{dI(p|\bar{\delta})}{dp} > \frac{dI(p|\delta)}{dp} \). This follows from the fact that

\[ \frac{dI(p)}{dp} = \frac{Ab_F R}{(1 + b_F + \frac{B(1-\delta_1E\delta_2)}{\delta_1E\delta_2} - pR)^2} \]

Thus, in the absence of a CRT

\[ \int_{p^*}^{p} I(p, \bar{\delta}) dG(p) > \int_{p^*}^{p} I(p, \delta) dG(p) \]

Thus, if this inequality is reversed, aggregate investment has increased and CRT is efficient.

\[ \blacksquare \]

Proof of Proposition 8

Recall, in the absence of a CRT market, \( M(p) = I(p | \delta)(1 - B_B) - A \). Thus, if there is no CRT aggregate losses on market debt are:

\[ \int_{p^*}^{p} (1 - p)[I(p | \bar{\delta})(1 - B_B) - A]dG > \int_{p^*}^{p} (1 - p)[I(p | \hat{\delta})(1 - B_B) - A]dG \]

Which follows from the fact that \( I(p | \bar{\delta}) > I(p | \hat{\delta}). \)

If a CRT is introduced, then market debt is \( I(p | \delta)(1 - B_B \frac{p}{p - r}) - A \). As the CRT is inefficient, total investment is lower. Further, \( \frac{p}{p - r} > 1 \) and increasing in \( p \). The result follows.

\[ \blacksquare \]
References


Figure 4