LOAN SALES AND THE COST OF BANK CAPITAL

by

George Pennacchi

7-87

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

Preliminary
Comments Welcome

Loan Sales and the Cost of Bank Capital

by

George Pennacchi*

February 1987

*Department of Finance, University of Pennsylvania, Philadelphia, PA 19104, (215) 898-6298. Comments by Franklin Allen, Mitchell Berlin, Michael Fishman, Joseph Haubrich, Robert Litzenberger, Krishna Ramaswamy, Michael Smirlock, and participants of the Finance Workshop at Penn and the conference on Asset Securitization at Northwestern are greatly appreciated. Funding provided by the Herbert V. Prochnow Educational Foundation of the Graduate School of Banking, Madison, Wisconsin.

Loan Sales and the Cost of Bank Capital

Abstract

This paper demonstrates that if banks are faced with significant competition for deposit financing, as well as regulatory constraints in the form of required capital and/or reserves, banks cannot be profitable solely by holding marketable assets. They must provide other services, such as information gathering and monitoring activities related to making loans. For a bank which originates loans, loan selling will likely provide a cheaper source of funds than traditional deposit or equity finance. However, the extent to which banks can sell loans is limited by the ability of the bank-loan buyer contract to overcome a moral hazard problem. The bank's choice of an optimal loan sales contract is analyzed with and without allowance for recourse.

1. Introduction

The practice of banks making loans and then selling them to other financial institutions and individuals has recently become quite popular. An important example of the development of loan selling has been the growth in banks' mortgage loans which are insured and pooled by the Government National Mortgage Association (GNMA) and then sold to secondary market investors. Within the last few years, there has also been a dramatic rise in the amount of other types of loans also being sold, especially by money center banks. Portions of commercial loans originated by these larger banks are being sold to pension funds, foreign banks, and smaller banks, a process also known as the selling of "sub-participations" in a loan. Banks' car loans have also been pooled and sold, and the selling of other types of consumer loans and credit is being initiated.[1] This expansion in sales of loans which were once thought of as "non-marketable" assets, a process which has come to be known as "asset securitization," may be signaling the start of a fundamental change in the way banks conduct business. The leading banks in loan selling operations are beginning to see themselves more as originators and distributors of loans, rather than institutions holding loans as assets.

Given the potentially large impact that loan sales may have in shaping the future of the commercial banking business, it appears quite relevant to ask what incentives exist for banks to sell loans. In this paper we show that loan sales may enable banks to tap a cheaper source of funds than traditional deposit or equity finance. However, the extent to which banks can sell loans is limited by the ability of the bank-loan buyer contract to overcome a moral hazard problem. This problem relates to the possible incentive for banks to inefficiently monitor and service loans after they have been sold. The capability of a bank to offer loan buyers an incentive efficient loan sales contract will determine the maximum amount of the bank's loan sales and therefore its profitability. We investigate the bank's optimal choice of a loan sales contract under different regulatory constraints.

The plan of the paper is the following. In section 2 we present a simple state-preference model of the banking firm. The bank must choose its optimal quantity of loans to originate as well as its optimal holdings of marketable assets. In addition the bank must decide whether to finance its asset holdings by issuing either deposits or equity. We first examine optimal portfolio and capital structure decisions when loan sales are not permitted, and second, when they are. In order to determine the equilibrium quantity of loan sales, the optimal bank-loan buyer contract must next be determined. In section 3 we analyze the optimal form that this contract takes when the level of loan monitoring by banks is unobservable and agents are risk neutral. Contracts are considered where the loan buyer has recourse to the bank for losses and also when no recourse is permitted. A conclusion follows in section 4.

2. A Model of the Banking Firm

This model concentrates on the loan making activities of banks. Making loans as well as possibly holding a collection of marketable assets is what we will loosely define as the "portfolio" services provided by banks; the process of channeling funds between savers and borrowers. It is assumed that in making loans, banks can expend real resources in gathering information on loan applicants and monitoring these loans so as to improve the return (quality) of these loans. The information gathering and monitoring functions of banks have been stressed in the papers by Campbell and Kracaw (1980), Diamond (1984), and Gorton and Haubrich (1986). Other services provided by banks, particularly convertibility, are ignored in this paper. As in Black (1970), Fama (1980), and Fischer (1982), it is assumed that no necessary connection exists between the portfolio and transactions services of banks, and hence they can be separated into different "departments" of the banking organization.[2]

2.1 Equilibrium with No Loan Sales

Let us consider a one period state-preference model where banks are assumed to have the opportunity of holding two different types of assets; marketable assets and non-marketable assets. Let M denote the value of marketable assets held by the bank. One can think of these marketable securities as money market instruments, i.e., short-term debt instruments such as Treasury bills. They are assumed to pay a certain rate of return (interest rate) of r_m over the period in each state of the world. Banks are assumed to be price takers for marketable assets, since the large number of traders in this market would suggest a high degree of competition.

The amount of bank funds invested in non-marketable assets will be denoted N. Non-marketable assets can be generally thought of as corporate, consumer, or agricultural loans. In terms of our model, the distinguishing feature of non-marketable assets is that a technology exists where a bank can improve its (uncertain) return on these assets by expending resources on information gathering and monitoring. Specifically, assume that banks can make unit loans such that for a given loan i, one dollar lent at the beginning of the period entitles the bank to an end of period cashflow equalling $x_i(s,a_i)$, where s indexes the state of nature at the end of the period and a_i is the level of "monitoring" chosen by the bank at the beginning of the period.[3] We assume a possibly infinite number of end of period states, so that s can be designated as $s \in [0, 1]$, i.e., real numbers on the unit closed interval.

A given loan's monitoring services, a_i, are produced via a constant returns to scale technology, using inputs such as computers, bank office buildings, labor of loan officers and assessors, etc. This implies that the cost function for monitoring services can be written as $c(a_i) = ca_i$ with c being a positive constant. In addition, it is reasonable to assume $x_i(s,a_i)$ is a monotonically increasing, concave function of a_i, $a_i \geq 0$.

Banks finance their asset purchases through issuing deposits or equity. Perfect competition is assumed to hold in both of these financial markets. Since this paper ignores transactions services of banks, one should think of these deposits as being in non-transactions accounts, such as Money Market Deposit Accounts, Certificates of Deposit, or other "purchased funds," and thus can be regarded simply as (short-term) debt instruments. In practice, non-transactions deposits have either zero or small reserve requirements. For example, Money Market Deposit Accounts have no required reserves while most Certificates of Deposit have a 3% requirement. For simplicity our model will assume one homogeneous class of deposits with required reserve ratio equal to ρ. Whether ρ is zero or small is not important in terms of our qualitative results, as will be shown later. Further, assume that all deposits are either legally or de facto insured by the government. Thus let r_d denote the interest rate on these deposit accounts for all states, s. The total level of the bank's deposit financing is given by D.

The amount of funds raised by the bank through equity issue is given by E. Assuming complete markets, let $p^e(s)$ denote the equilibrium price (density) paid by equityholders at the beginning of the period for a security that pays one dollar in state s at the end of the period, where the return on this security is treated, for personal tax purposes, as equity. For example, the market price paid by equityholders for an equity share which would pay one dollar for all states, $s \in [0, 1]$, is given by;

(1) $$\int_0^1 p^e(s)ds \equiv \frac{1}{1 + r_e}$$

r_e can be thought of as the certainty equivalent cost of equity financing.

The government also plays a role in this model via taxation of bank profits, the regulation of bank capital, and the provision of deposit insurance. Let t be the proportional corporate income tax rate and let ζ be the maximum deposit (debt)-equity ratio allowed the bank, i.e., ζ represents a "capital adequacy" standard.

The method used by the government to price deposit insurance must be specified. Distortions created by the mis-pricing of deposit insurance have been studied in a number of papers, although it is a controversial question as to whether deposit insurance is actually over or under-priced.[4] Because we choose to focus on somewhat different topics, and because mis-priced insurance is not likely to lead to qualitatively different results, the simplifying assumption that deposit insurance is priced fairly is made.[5] It is shown in the Appendix that if the bank's objective is to maximize the after tax gain to shareholders' equity, the fair insurance pricing assumption leads to the following form for the objective function;

(2) $$\max_{\{N,(a_i),M,D,E\}} \{(1+r_e) \sum_{i=1}^{N} \int_0^1 p^e(s)x_i(s,a_i)ds - N - c \sum_{i=1}^{N} a_i + r_m M - r_d D\}(1-t) - r_e E$$

subject to the constraints;

(3) $$N + M + \rho D \le D + E$$ (Financing Constraint)

(4) $$D \leq \zeta E$$ (Capital Constraint)

In expression (2) we choose the bank loans i, i = 1, . . ., N to be ordered from the highest valued loan to the lowest.[6] Thus loan i = N is the "marginal" loan made by the bank, i.e., the loan which the bank is (approximately) indifferent to making.

Differentiating with respect to each of the bank's choice variables, this leads to the first order Kuhn-Tucker conditions;[7]

(5) $$\left[(1 + r_e)\int_0^1 p^e(s)x_N(s, a_N)ds - 1 - ca_N\right](1 - t) - \lambda_1 = 0$$

(6) $$\left\{(1 + r_e)\int_0^1 p^e(s) \frac{\partial x_i}{\partial a_i} (s,a_i)ds - c\right\}a_i = 0$$

(7) $$\{r_m(1 - t) - \lambda_1\}M = 0$$

(8) $$\{-r_d(1 - t) + \lambda_1(1 - \rho) - \lambda_2\}D = 0$$

(9) $$\{-r_e + \lambda_1 + \zeta\lambda_2\}E = 0$$

where λ_1 and λ_2 are the multipliers associated with constraints (3) and (4), respectively, and the expressions in brackets are all non-positive.

Let us examine the bank's optimal capital structure first. Conditions (8) and (9) determine the bank's optimal choice of debt versus equity. The relevant question is whether the bank's capital constraint (4) is binding, or whether the bank will optimally prefer more equity than that required by (4). Quite simply, banks will choose to be at their maximum debt-equity ratios, ζ, if;

(10) $$r_d \frac{(1 - t)}{(1 - \rho)} < r_e .$$

Now if this inequality is reversed, banks will prefer to finance their assets entirely through issuing equity. However, there is good reason to believe that (10) does hold for banking firms, that their capital adequacy constraint is binding, and we sketch an argument for this below. Let us first consider the case of zero required reserves, i.e., $\rho = 0$.

Although the well-known work of Miller (1977) presents a general equilibrium rationale for $r_d(1 - t) = r_e$ such that individual firms are indifferent between the choice of debt versus equity finance (though the aggregate debt-equity ratio is determinate), theoretical extensions to Miller's model tend to restore the existence of an optimal debt-equity ratio at the individual firm level.[8]

Using a state-preference framework similar to that of this paper, DeAngelo and Masulis (1980) show that if some firms in the economy have alternative methods to reduce their tax liabilities, such as depreciation cost allowances or investment tax credits, an equilibrium will likely exist where $r_d(1 - t) > r_e$. Barnea, Haugen, and Senbet (1981) also examine how the Miller equilibrium might be modified when agency costs of debt financing are considered. They show that if agency costs to debt issue are an increasing function of a firm's debt level, an equilibrium will obtain in which firms with low firm specific agency costs of debt have high optimal debt-equity ratios whereas firms with high agency costs will choose low leverage. Because of deposit insurance, it is quite reasonable to believe that banks' agency costs of deposit or debt finance are very low, or even zero, and hence their optimal leverage is very high. Deposit insurance transfers the problems of moral hazard or agency cost from depositors to the deposit insuring agency, again implying that capital constraints will be binding.[9] Further, Marcus (1983) presents relevant empirical evidence consistent with the hypothesis that regulatory costs restrict banks' optimal debt-equity levels. There is also substantial casual evidence that capital constraints are binding since many bankers see increases in capital requirements as costly.[10] Thus even with effective required reserves being positive but sufficiently small, there seems much justification in believing that at the margin deposit finance is less expensive, overall, than equity finance, so that condition (10) holds.

Given this binding capital constraint, the bank's optimal level of non-marketable assets can now be determined. From the first order conditions (5), (6), (8), and (9) we have that the bank's choice of loan volume, N, will satisfy;

$$(11)\quad (1 + r_e)\int_0^1 p^e(s)x_N(s,a_N^*)ds - 1 = ca_N^* + \left[\frac{r_e}{1 - t} + \zeta r_d\right]/(1 + \zeta(1 - \rho)) \ .$$

The left-hand side of (11) is the value of contingent interest income received from the marginal one dollar bank loan when the level of monitoring, a_N^*, is optimally chosen according to condition (6). The right-hand side of (11) is simply the cost of making this marginal one dollar loan. The first term on the right is the additional resource cost of optimally monitoring while the last term is the bank's weighted marginal cost of capital, reflecting the fact that since the bank's capital requirement is binding, it must issue both new equity and deposits in order to expand loans.

Consider next the bank's choice of marketable assets, which we interpret to be money market instruments such as Treasury bills, commercial paper, banker's acceptances, and negotiable certificates of deposit. These securities can be held directly by individuals or indirectly through money market mutual funds, but in either case their personal income tax treatment is the same as that of a short-term debt instrument or bank deposit. Now if bank assets each have the same personal tax treatment, risk, and liquidity, we can, following Fama (1985), treat them as perfect substitutes. Therefore, in a competitive equilibrium they must have the same yield;

$$r_d = r_m . \tag{12}$$

What does this imply concerning banks' choice of marketable assets to hold in their portfolio? Using conditions (7) and (10) one obtains;

$$r_m(1 - t) - \frac{r_e}{1 + \zeta(1 - \rho)} - \frac{\zeta r_m(1 - t)}{1 + \zeta(1 - \rho)} < 0 \tag{13}$$

and hence M = 0. This non-trivial result says that if we assume marketable assets are priced competitively, banks, relative to money market funds, are inefficient intermediaries through which investors could hold marketable assets.[11] Because banks are required to issue a proportion ζ of equity for each dollar deposit and are subject to corporate taxes when mutual funds are not, even if required reserves equal zero, banks' asset portfolios must yield more than just the return on marketable assets. Discounted by the weighted cost of banks' capital, marketable assets produce negative net present value for banking firms.

Therefore banks must find other types of investments, such as making loans, which are activities not in direct competition with money market funds or individual investors.[12] The next section shows how loan selling is likely to be a profitable arrangement for banks which originate loans.

2.2 Equilibrium with Loan Sales

Consider a bank selling a claim on the return of a loan it has originated. It is assumed that the interest on the loan buyer's funds used to buy this claim is taxed in the same manner as the interest on a standard in more detail in the next section, here we simply assume that a share, b_i, of the return from each loan i is sold to a loan buyer. Thus if loan i originated by the bank requires a dollar of initial financing and returns $x_i(s,a_i)$ dollars if state s occurs and the bank monitors at level a_i, loan buyers receive $b_i x_i(s,a_i)$ while the bank receives $(1 - b_i)x_i(s,a_i)$.

The value of the loan share sold by the bank is specified in a manner which is similar in spirit to that of Litzenberger and Van Horne (1978) and DeAngelo and Masulis (1980). A second set of primitive securities, $p^d(s)$, $s \in [0, 1]$, whose return is treated, for personal tax purposes, as debt is assumed to exist.[13] These securities denote the equilibrium prices paid by debtholders for one dollar debt returns in each state of nature. One might think of these securities as "junk" bonds. By the definition of r_d, it must be that;

$$(14) \qquad \int_0^1 p^d(s)ds \equiv \frac{1}{1 + r_d} .$$

The amount paid by loan buyers for a share, b_i, of loan i is then;

$$(15) \qquad \bar{b}_i \equiv b_i \int_0^1 p^d(s)x_i(s,a_i)ds$$

Suppose banks finance a given proportion, $\bar{b}$, for each dollar loan made through loan sales, i.e., $\bar{b}_i = \bar{b}$, i= 1, . . ., N. Assuming fairly priced deposit insurance, one obtains an objective function for the bank;

$$(16) \qquad \max_{\{N,(a_i),D,E\}} \{(1+r_e)\int_0^1 p^e(s) \sum_{i=1}^N (1-b_i)x_i(s,a_i)ds-(1-\bar{b})N-c \sum_{i=1}^N a_i-r_dD\}(1-t)-r_eE$$

subject to the same constraints (3) and (4) as before, expect that now $(1 - \bar{b})N$ replaces N in (3).

With the opportunity of selling loans, the bank's optimization problem now involves the following equilibrium condition regarding its choice of loans originated, N;

(17)
$$(1 + r_e)\int_0^1 p^e(s)x_N(s,a_N)ds - 1 = ca_N + r_I$$
$$- b_N\left[(1 + r_I)\int_0^1 p^d(s)x_N(s,a_N)ds - (1 + r_e)\int_0^1 p^e(s)x_N(s,a_N)ds\right]$$

where $r_I \equiv \left[r_e/(1 - t) + \zeta r_m\right]/\left[1 + \zeta(1 - \rho)\right]$ is the bank's weighted marginal cost of internal financing, which was the bank's marginal cost of capital for the no loan sales case. Comparing the form of equation (17) with the analogous condition (11) for the no loan sales case, we see that the marginal cost of originating a loan, which is the right-hand side of (17) differs from that of (11) because of an additional final term. This term represents a possible savings to the bank in its marginal cost of capital due to raising funds via loan sales.

Indeed, under reasonable circumstances, it is likely that this last term on the right-hand side of (17) will be less than zero. To see this, note that equations (1) and (14) can be combined to obtain;

(18)
$$\int_0^1 p^d(s)ds/\int_0^1 p^e(s)ds = (1 + r_e)/(1 + r_d)$$

This implies that the ratio of the "averages" of the primitive security prices for debt and equity equal the ratio of the certainty equivalent rates of return on equity versus debt. An additional assumption which is sufficient, though not necessary, for loan sales to lower the marginal cost of capital would be that the debt-equity security price ratio be uniform across states;

(19) $$p^d(s)/p^e(s) = (1 + r_e)/(1 + r_d) \quad \text{for all } s.$$

DeAngelo and Masulis (1980) show that investor risk neutrality would imply this uniformity. Employing this stronger assumption, the last term in equation (17) can be re-written;

(20) $$- b_N[(1 + r_e)\int_0^1 p^e(s)x_N(s,a_N)ds \frac{(r_I - r_d)}{(1 + r_d)}] < 0$$

since $r_I > r_d$ from condition (10).

Ideally, we would like to have a general equilibrium model determining the relative price of the two sets of securities for each state in order to exactly determine whether the sale of a given loan lowers the cost of bank financing. However, developing a model of this sort would lead us far from the focus of this paper. Still, from an intuitive level, the loan sales equilibrium condition (17) indicates that if the price paid by loan buyers for the loan sale debt claim sufficiently reflects the higher average relative price paid for debt securities over equity securities, then the bank can lower its cost of financing by selling loans. Consider the simple case of complete certainty, i.e., a single end of period state of the world. By selling loans banks can raise funds at the same cost, $r_d = r_m$, as deposits. However, since the funds acquired through loan sales do not appear as a larger level of deposits on the balance sheet of the bank, the bank will not be required to issue more relatively expensive equity in order to stay within its capital adequacy constraint, or hold non-interest paying reserves against these funds. Therefore in this certainty case, loan sales are definitely a method for banks to reduce costs stemming from capital controls and required reserves. Ceteris paribus, with a lower cost of capital, it will be profitable for banks to expand their loan originating and monitoring activities, leading to a lower competitive interest rate on loans. From a macro-economic perspective, an economy wide increase in the proportion of loans sold would decrease the demand for high powered money, likely raising nominal measures of output.

Thus far, we have neglected a rather important issue relating to the levels of monitoring that the bank chooses when it decides to sell loans. If the bank's monitoring levels, a_i, i = 1, . . ., N, are observable by loan buyers, enabling the bank to commit to given levels of monitoring, then loan sales can, indeed, lead to greater bank profits and greater numbers of loans being originated. To see this, suppose banks can commit to the same levels of monitoring as in the no loan sales case, given by condition (6), i.e., $a_i = a_i^*$. Then a comparison of the marginal loan equilibrium conditions (11) and (17) will imply a lower total marginal cost of an additional one dollar loan, given the negativity of the last term on the right hand side of (17). Furthermore, when monitoring is observable, if it is profitable to sell any share of a loan, it must be even more profitable to sell the entire loan.

However, when loan monitoring is unobservable by loan buyers, a potential moral hazard problem comes into the picture which can limit the proportion of loans sold. The problem arises from the bank's diminished incentive to monitor, since its monitoring level will now satisfy;

$$(1 + r_e)(1 - b_i)\int_0^1 p^e(s) \frac{\partial x_i(s,a_i)}{\partial a_i} ds = c \tag{21}$$

With the marginal benefit from monitoring being diminished by the factor $(1 - b_i)$, monitoring will be less than in the no loan sales case, and this will tend to lower the share price paid by loan buyers, the greater the share of the loan that is sold. However, by optimally structuring the bank-loan buyer contract, this moral hazard problem can be reduced. Determining this optimal contract will help to explain the proportion of loans that can be sold.

3. Bank-Loan Buyer Contract Choice

A bank's decision to originate and monitor a given loan can be made independently of the same decisions for other loans, as can be verified from the banks' objective function (16) and conditions (17) and (21). Therefore a separate objective function for each given loan can be written down. Taking condition (16) for the case of N = 1, multiplying by (1 - t), and substituting in the equilibrium condition $D = \zeta E$, one obtains the individual loan objective function;

$$(22) \qquad (1 + r_e)\int_0^1 p^e(s)(1 - b)x(s,a)ds - ca - (1 + r_I)I$$

where $I \equiv (1 - \rho)D + E = 1 - \bar{b}$ is the amount of internal financing used in originating the one dollar loan.

In order to add more structure to our problem and keep the analysis tractable, the case in which security valuation by shareholders and loan buyers reflects risk neutrality is considered. In this case investors will only be concerned with the expected return on their contingent claim to the loan. In addition, we make the following assumptions regarding the return distribution of the bank loan, the cost of monitoring, and observability.

A.1. The stochastic return on the loan, x, has a distribution such that $x \in [0, L]$ where L is the promised end of period payment on the loan. The bank can alter the loan's return distribution by monitoring, such that the probability density function of the loan's return has the form f(x,a). It is assumed, as in Hart and Holmstrom (1986), that the loan's distribution function, F(x,a), satisfies the Convexity of Distribution Function Condition (CDFC);

$$(23) \quad F\left(x, \lambda a + (1 - \lambda)a'\right) \le \lambda F(x,a) + (1 - \lambda)F(x,a'), \ \forall a, a'; \ \lambda \in (0, 1)$$

A.2. Let the bank's cost of monitoring a loan be given by c(a) where $c'(a) > 0$ and $c''(a) \ge 0$.

A.3. Bank loan monitoring is unobservable by loan buyers. They can, however, observe the loan's actual return, and hence their share of the loan's return may be contingent on the loan's actual return, i.e., $b = b(x)$.

The bank's problem of choosing the optimal loan sales contract and level of monitoring can then be written as;

$$(24) \quad \underset{\{b(x),a\}}{\text{Max}} \int_0^L \left(1 - b(x)\right) x dF(x,a) - c(a) - (1 + r_I)I$$

subject to;

$$(25) \quad \int_0^L b(x) x dF(x,a)/(1 + r_d) + I = 1 \qquad \text{(Financing Constraint)}$$

$$(26) \quad \int_0^L [1 - b(x)]xdF(x,a) - c(a) \geq \int_0^L [1 - b(x)]xdF(x,a') - c(a') \ \forall a' \neq a$$

(Incentive Compatibility Constraint)

Substituting out for I in (24) using the financing constraint, it is straightforward to show that the objective function can be converted to;

$$(27) \qquad \underset{\{b(x),a\}}{\text{Max}} \int_0^L \ [1 + \theta b(x)]xdF(x,a) - c(a) - (1 + r_I)$$

where $\theta = (r_I - r_d)/(1 + r_d)$ is the present value of savings on the cost of capital by financing through loan sales rather than internal funds. Also, if condition (23) holds, Hart and Holmstrom (1986) show that the incentive compatibility constraint, (26), can be converted into the more convenient form;

$$(28) \qquad \int_0^L \ [1 - b(x)]xdF_a(x,a) = c'(a)$$

We are now prepared to consider a variety of contractual arrangements between the bank and the loan buyer.

3.1 Loan Sales without Recourse

We will define a bank-loan buyer contract with no recourse as one in which the bank cannot pledge outside assets as a potential payment to the loan buyer. Only the proceeds of the loan return are permitted to be split between the bank and loan buyer. As will be discussed in more detail in the next section, the Federal Reserve places restrictions on the type of recourse on losses that banks can offer loan buyers, and in practice, many commercial loan sales are made without recourse. Therefore it is of interest to consider no recourse loan sales. The no recourse restriction on the loan sales contract takes the form:

(29) $$b(x) \leq 1 \quad \text{for all } x \ .$$

Therefore (27), (28), and (29) characterize the bank's problem of selecting the optimal no recourse contract. Let λ be the Lagrange multiplier for constraint (28) and $\mu(x)$ be the multipliers for the inequalities in (29).

The first order condition with respect to the bank's choice of b(x), for a given x, is;

(30) $$\{\theta x f(x,a) - \lambda x f_a(x,a) - \mu(x)\} b(x) = 0 \ .$$

where the expression in brackets must be non-positive. The first order condition regarding the bank's level of monitoring is then;

(31) $$\frac{(1 + r_I)}{(1 + r_d)} \int_0^L b(x) x dF_a(x,a) + \lambda \left[\int_0^L \{1 - b(x)\} x dF_{aa}(x,a) - c''(a) \right] = 0$$

Before attempting to analyze equations (30) and (31), let us consider the characteristics of the probability density function, f(x,a), for a typical bank loan, as this will prove insightful in interpreting these optimality conditions. Let us assume a bank loan is made to an otherwise all equity financed firm which invests its funds in assets (projects) with an uncertain return. If V is the value of this firm's assets when its loan with promised payment L becomes due, then at maturity the value of this bank loan will be;

(32) $$x = \min[L, V]$$

A reasonable assumption concerning the range of the distribution of V is that it is bounded below at zero. In addition we assume that the bank's monitoring level, a, affects the form of the firm's asset density function such that a lower level of 'a' implies a "fatter" lower tail of the density function of V. Figure 1 gives a plausible form for this probability density function of V. $g(V,a_1)$ is the density if the firm is monitored by the bank at level a_1 while $g(V,a_0)$ is the density if the firm is monitored at level a_0, where $a_1 > a_0$.

Given the density function for V, the density function for the loan return, x, is determined. The loan return density when the bank monitors at level a_1, $f(x,a_1)$ is simply equal to $g(V,a_1)$ for $V < L$, and with all the probability mass of $g(V,a_1)$ for $V \geq L$ "piled" together at point L. Thus the value of $f(L,a_1)$ is a Dirac delta function spike with area equal to the prob $(V \geq L)$. $f(x,a_0)$ over the range [0, L] will bear a similar relationship to $g(V,a_0)$.

Note in Figure 1 that if the promised loan payment, L, is not too large relative to the density of V, then $g(V,a_0) > g(V,a_1)$ and hence $f(x,a_0) > f(x,a_1)$ for all V and x less than L. In other words, if the promised loan payment is sufficiently in the lower tail of the firm's asset return, then over the range 0 to L the density is a decreasing function of the bank's monitoring level, i.e., less monitoring makes the tail "fatter."[14] The casual observation that banks rarely make commercial loans carrying exorbitant interest rates, e.g., 20 points above prime, lends support to the proposition that L is typically in the lower tail of the firm's asset distribution, where $g_a(V,a) < 0$.

Now assuming as in Figure 1, that $f_a(x,a) < 0$ for all $x < L$, we see from condition (30) that the expression within brackets must be non-negative for all $x < L$ such that

$$\mu(x) = \theta x f(x,a) - \lambda x f_a(x,a) \geq 0 \qquad \forall\ x < L \tag{33}$$

Therefore $b(x) = 1$ for all $x < L$, i.e., the optimal loan sale contract gives the loan buyer the entire loan return whenever a loan default occurs. Hence the bank will receive a return from the loan only when the loan does not default, since only when $x = L$ will $f_a(L, a)$ be positive. Only in this case will $\mu(L) = 0$ so that $b(L) < 1$ and;

$$\lambda = \theta \frac{f(L,a)}{f_a(L,a)} . \tag{34}$$

Thus our assumptions on the loan distribution and preferences lead to a unique piece-wise linear optimal sharing rule that looks very similar to the loan buyer having a debt position and the bank an equity position in the loan.[15] The contract is characterized by penalizing the bank if low loan outcomes occur, and rewarding the bank if high loan outcomes (no default) occurs. Giving the bank a disproportionate share of the risk allows the bank to reap a disproportionate share of the gains from monitoring, enabling a greater amount of the loan to be sold while maintaining monitoring incentive efficiency.

There is some evidence that actual non-recourse loan sales contracts follow this principle of giving the selling bank a disproportionate share of the loan's risk. One example is the case of money center banks which have begun the practice of selling short-term "strips" of loans. This involves banks selling claims on the earlier payments of a multi-payment loan, i.e., payments with maturities generally in the range of 30 days to three months, and hence with relatively low default risk. The selling bank retains claim to the later more risky payments from the loan, generally with maturities from one to seven years, and thus preserves much of its incentive to monitor.[16] Melvin (1986, p. 41) cites another example regarding the case of selling promised payments from a pool of credit card receivables in which the selling bank retains an equity position in the pool equal to twice the historical default level of the receivables.

Let us now examine the equilibrium level of bank monitoring that will result under the optimal loan sales contract. Note that the terms in brackets in equation (31) is just the second order condition regarding the bank's optimal monitoring choice, which is assumed to hold and therefore is negative. This implies the loan buyer's expected benefit from greater monitoring is positive;

$$\int_0^L b(x)x dF_a(x,a) > 0 \tag{35}$$

Using the result that $b(x) = 1$ for $f_a(x,a) < 0$ (whenever $x < L$) and $b(x) < 1$ for $f_a(L,a) > 0$, we also have;

$$\int_0^L b(x)x dF_a(x,a) < \int_0^L x dF_a(x,a) \tag{36}$$

Rearranging the bank's incentive compatibility condition (28), we can further show that

$$\int_0^L x dF_a(x,a) = c'(a) + \int_0^L b(x)x dF_a(x,a) > c'(a) \tag{37}$$

Since the left-hand side of (37) is the expected marginal return on the loan from additional monitoring which exceeds the marginal cost to greater monitoring, we conclude that the equilibrium level of monitoring by the bank will be less than the most economically efficient (first-best) level. We now turn to the case of the optimal contract design of loans sold with recourse.

3.2 Loan Sales with Recourse

One simple contractual arrangement which could provide a first-best solution to the above problem (27)-(28), would be for the bank to simply guarantee the loan buyer a rate of return of r_d regardless of the actual loan payoff. This contract, similar to the bank-depositor contract in Diamond (1984), would be feasible if the bank was able to maintain sufficient asset (loan) diversification, such that the probability of the bank's failure is negligible. Giving the loan buyer recourse to claims on other bank assets would conceivably allow the bank to sell the entire loan, and still retain the incentive to monitor at the economically efficient level.

However, the Federal Reserve has sought to place restrictions on direct guarantees on loan sales.[17] With only a few exceptions, guarantees by banks to reimburse loan buyers for loan losses, even if a ceiling on the amount of the bank's reimbursement is made, would lead the Fed to treat the bank's proceeds from a loan sale as a "deposit" subject to inclusion in calculations of required capital and possibly subject to required reserves. Federal Reserve proposals do allow for loans to be sold with recourse in the following manner: If a bank agrees to guarantee a given <u>percentage</u> of loan sales losses, say ℓ, where ℓ is less than 75%, then this bank would be permitted to classify only a proportion $(1 - \ell)$ of the proceeds from the loan buyer as a loan sale. The other proportion, ℓ, of the proceeds must be classified as a deposit, again subject to required capital. However, the Federal Reserve has stated that as long as the percent losses guaranteed, ℓ, is less than 75%, no required reserves need be held against the proceeds of the loan sale.

Under this arrangement, where the loan buyer's payment is $b(x)x = bx + b\ell(L - x)$, we might ask what is the bank's optimal choice of b and ℓ? For this type of recourse loan, the bank's problem is;

$$(38) \qquad \max_{\{b,\ell,a\}} \; \bar{x}(a) - b\left[\bar{x}(a) + \ell\left(L - \bar{x}(a)\right)\right] - c(a) - (1 + r_I)I$$

subject to;

(39) $$b\left[\bar{x}(a) + \ell\left(L - \bar{x}(a)\right)\right]\left(\frac{1 - \ell}{1 + r_d} + \frac{\ell}{1 + r_c}\right) + I = 1$$

(40) $$\left[1 - b(1 - \ell)\right]\bar{x}_a = c'(a)$$

(41) $$0 \le \ell \le \bar{\ell} = .75$$

(42) $$0 \le b \le 1$$

where $\bar{x}(a) = \int_0^L xdF(x,a)$, $\bar{x}_a = \int_0^L xdF_a(x,a)$, and r_c denotes the cost of bank funds from loan sales which are subject to capital constraints but not required reserves. r_c will be the cost of capital on the proportion ℓ of loan sale proceeds when the proportion of loan losses guaranteed is less than 75%. Therefore r_c equals the expression for r_I but where required reserves, ρ, has been set equal to zero, and hence $r_c \le r_I$.

Assuming the incentive compatibility constraint (40) is binding in equilibrium, it is clear from (41) and (42) that the equilibrium level of bank monitoring, a, will be less than the economically efficient (first-best) level. Interpreting the Kuhn-Tucker conditions from (38)-(42), it is straightforward to show that possible optima exist only for two sets of (b, ℓ) combinations; where constraints (40) and (41) bind or where constraints (40) and (42) bind. Therefore banks will always choose a positive level of loss guarantees under these regulations. Depending on the magnitude of the parameters of the model, the bank will either sell the entire loan with $\ell < \bar{\ell}$ or choose to sell somewhat less than the entire loan, with $\ell = \bar{\ell}$.

4. Conclusion

If banks are faced with significant competition for deposit financing, as well as regulatory constraints in the form of required capital and/or reserves, this paper has shown that banks cannot be profitable by avoiding the type of "passive" portfolio services offered by mutual funds. Banks' economic viability must derive from other services, such as information gathering and monitoring activities related to making loans. Given that banks choose to originate loans, loan sales will likely reduce banks' cost of financing.

However, a bank's ability to sell loans depends on loan buyers' perception of the bank's level of monitoring of these loans, which in turn is determined by the bank's incentive to monitor. By designing the loan sales contract in a way that gives the bank a disproportionate share of the gains to monitoring, a greater share of the loan can be sold and hence a greater level of bank profits can be attained. While the optimal loan sales contract is designed to preserve the bank's incentive to monitor, our model showed that for loans sold both with and without recourse, the equilibrium level of monitoring will be less that if the loan was not sold by the bank.

There are other issues concerning loan sales which this paper has not explored but which may be productive areas of future research. Our treatment of the optimal loan sales contract was essentially that of a one-period problem. In practice, a bank may acquire over many "periods" a reputation as an efficient monitor of loans sold. This could result in the bank being able to sell an even greater share of loans than our model would predict. The cost to losing this reputation may then act as a deterrent to the bank choosing a less efficient level of monitoring.

The effect that greater loan sales have on overall bank risk has also not been adequately addressed. Given that the optimal loan sales contract attempts to give the bank a large share of the gains to monitoring, which generally implies giving the bank a greater degree of the loan's risk, it might appear that loan sales would increase the volatility of the bank's asset portfolio. However, for a given bank capital structure, it is not clear that the benefits of asset diversification, deriving from a greater number of loans originated when loan selling occurs, might not outweigh the higher risk incurred on each individual loan. It may be unwise for regulators to unconditionally discourage loan sales.

Appendix

Below is a derivation of the bank's objective function when banks maximize the after-tax rate of return to shareholders and deposit insurance is fairly priced. Using the notation in the text, the bank's end of period after tax asset value, when state s occurs, is;

$$\text{(I)} \quad \{\sum_{i=1}^{N} x_i(s,a_i) - c \sum_{i=1}^{N} a_i - N + r_m M - r_d D\}(1 - t) + N + M + \rho D + r_d D$$

Letting ϕ be the premium charged for deposit insurance, the payoff to equityholders when the bank is solvent is;

$$\text{(II)} \quad \{\sum_{i=1}^{N} x_i(s,a_i) - c \sum_{i=1}^{N} a_i - N + r_m M - r_d D\}(1 - t) + E - \phi D \equiv W(s) - \phi D$$

Of course equityholders receive nothing when bankruptcy occurs. The end of period payoff to the deposit insurer is ϕD when the bank is solvent and is W(s) when the bank is insolvent. The insurer receiving W(s) implicitly assumes that the value of the bank's tax shield is preserved when the bank fails. This is not unrealistic since Kane (1985, p.38) points out that when a failed bank is mergered with an acquiring bank, the failed bank's losses can be used to reduce the acquiring bank's tax liability.

Letting S_0 denote the set of solvency states, and S_1 denote the set of insolvency states, a fair deposit insurance premium, where "fair" is used in the sense of an equivalent valuation in terms of primitive equity security prices, is ϕ such that;

$$\text{(III)} \qquad \phi D \int_{S_0} p^e(s)ds + \int_{S_1} p^e(s)W(s)ds = 0$$

Substituting this value for ϕ in (II) and taking the shareholders' present value of (II) over all solvency states, S_0, one obtains;

$$\text{(IV)}\quad \left\{(1+r_e)\int_0^1 p^e(s) \sum_{i=1}^{N} x_i(s,a_i)ds - c \sum_{i=1}^{N} a_i - N + r_m M - r_d D\right\} \frac{(1-t)}{(1+r_e)} + \frac{E}{1+r_e}$$

If the bank is assumed to maximize the difference between the present value of equityholders payment (IV) and the amount of equity that must initially be raised, E, then by subtracting E from (IV) and multiplying by $(1 + r_e)$ one obtains equation (2) in the text.

Footnotes

[1]See Salem (1985) and Pavel (1986) for a description of current developments regarding bank loan selling and the more general phenomenon known as "asset securitization." Also relevant is: NY Times, January 20, 1986, "Loan Sales Market Swelling," and February 11, 1985, "Repackaging of Car Loans is Increasing."

[2]Extending the model to allow for a transactions technology and transactions deposit accounts produces no substantive changes in the model's results concerning loan sales.

[3]While this effort expended by the bank for a given loan, a_i, is referred to as the bank's level of monitoring activities, this effort could also be interpreted to include the bank's information gathering and credit checking activities necessary to select a better quality loan from a pool of applicants. Therefore our reference to a moral hazard problem of inefficient monitoring caused by the unobservability of the bank's effort by loan buyers, which is explained in section 2.2 and treated in section 3, could be interpreted to also refer to an adverse selection problem of inefficient information gathering by the bank.

[4]In a similar state-preference framework, Dothan and Williams (1980) analyze distortions arising from the mis-pricing of deposit insurance. Pennacchi (1987) examines whether deposit insurance provided by the FDIC is generally over or under priced.

[5]Because of the presence of a capital adequacy constraint in our model, unlike Dothan and Williams (1980), the incentive for banks to pursue greater risk normally associated with fixed rate or subsidized deposit insurance is reduced, because greater risk cannot be obtained via higher leverage. Banks may increase their risk by less stringent monitoring of loans, but in our model, this would also result in a decrease in the loan's value. If by lowering monitoring, the loss to the bank from a fall in loan value exceeds the bank's gain from a greater value of deposit insurance, banks may choose not to pursue a higher risk strategy, but instead follow firm value maximizing behavior as in the case of the fair pricing of insurance.

[6]This is assuming that each a_i is chosen optimally according to condition (6).

[7]Condition (5) is obtained by differentiating with respect to N but ignoring the integer constraint on loans.

[8]Miller (1977) reasons that while returns from equity finance are more heavily taxed than debt at the corporate level, for most individuals equity returns are taxed at a lower rate than debt at the personal income tax level. Under certain conditions he shows that an equilibrium with "tax clienteles" could exist where high tax individuals hold equities and low tax individuals hold debt. While the aggregate corporate supplies of debt and equity are determinate, such that $r_d(1 - t) = r_e$, individual firms will be indifferent to their debt-equity mix. However, there is ample empirical

Also note that with tax law changes resulting in a 34% corporate tax rate and only a 28% marginal tax rate for high income individuals, it could be argued that a Miller equilibrium is now impossible.

[9]Buser, Chen, and Kane (1981) present a model which illustrates this point.

[10]See, for example, "Capital-Ratio Rise Sours Bank Growth," The Wall Street Journal, December 2, 1985.

[11]Fama (1985) gives empirical evidence and a theoretical explanation similar to that of this paper to show why Certificates of Deposit will have a yield nearly identical to similar money market instruments, even though required reserves are imposed on C.D.'s but not on other money market assets. Because the "reserve tax" falls not on depositors but on the bank, Fama reaches the same conclusion as this paper, that banks issuing C.D.'s will choose to "hold no open-market securities" (marketable assets). See Fama (1985, p. 34). Note, however, that our model predicts banks would hold no marketable assets even if required reserves were zero, as banks are still constrained to hold relatively expensive capital.

[12]Of course, there are other activities which banks can find profitable, such as the provision of transactions services (e.g., check writing) since money market funds provide only limited competition in these services.

[13]In a Miller (1977) type of world with investors in different tax clienteles, the ratio of the rates of return for a primitive debt and equity security would equal the ratio of the complements of the personal tax rates for equity and debt for the "marginal" investor. See Litzenberger and Van Horne (1978), page 739.

[14]One can think of the function of the bank's monitoring to be that of limiting the risk of the borrowing firm's projects (assets). The bank, by reducing the "fatness" of the tail of the firm's asset distribution, is improving the expected return on its loan.

[15]The contract is not exactly debt-equity division of the loan return. Note that for small loan default, i.e., $x = L - \varepsilon$, where ε is a small positive quantity, the loan buyers could receive a total return greater than their return when the loan did not default at all, since $b(x) = 1$ for $x < L$ and $b(x) < 1$ for $x = L$.

The framework of Holmstrom (1979) can be used to derive contract optimality conditions under an alternative assumption that bank-loan buyer asset choice displays risk-aversion.

[16]See "Major New York Banks Initiate Tactic of Selling Short-Term 'Strips' of Loans," The Wall Street Journal, January 23, 1986. Merton (1974) shows that the risk premium on a promised corporate payment rises as the time until payment is received increases.

[17]See Federal Reserve regulations 12 CFR Part 204 Regulation D; Docket No. R-0571 and the instructions for filing Reports of Condition and Income, and also the explanations of these regulations in Pavel (1986). While the

sales, they have not acted to restrict loan sales which are guaranteed by third party insurance companies, even if the bank and insurance company negotiate an agreement which obligates the bank to re-imburse ex-post the insurance company for any payments it must make to the loan buyer. Under these circumstances, third party insurance of loan sales appears to be the optimal arrangement from the bank and loan buyer's point of view. However, it is reasonable to believe that the existence of this loophole will be short-lived.

References

Barnea, A., R. Haugen and L. Senbet, 1981, "An Equilibrium Analysis of Debt Financing under Costly Tax Arbitrage and Agency Problems," Journal of Finance 6, June, 569-581.

Black, F., 1970, "Banking and Interest Rates in a World Without Money," Journal of Bank Research 1, Autumn, 8-20.

Buser, S., A. Chen and E. Kane, 1981, "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital," Journal of Finance 35, March, 51-60.

Campbell, T. and W. Kracaw, 1980, "Information Production, Market Signalling and the Theory of Financial Intermediation," Journal of Finance 35, September, 863-881.

DeAngelo, H. and R. Masulis, 1980, "Optimal Capital Structure under Corporate and Personal Taxation," Journal of Financial Economics 8, March, 3-29.

Diamond, D., 1984, "Financial Intermediation and Delegated Monitoring," Review of Economic Studies 51.

Dothan, U. and J. Williams, 1980, "Banks, Bankruptcy, and Public Regulation," Journal of Banking and Finance 4, 65-87.

Fama, E., 1985, "What's Different About Banks?" Journal of Monetary Economics 15, 29-39.

Fama, E., 1980, "Banking in the Theory of Finance," Journal of Monetary Economics 4, 39-57.

Fischer, S., 1982, "A Framework for Monetary and Banking Analysis," National Bureau of Economic Research, Working Paper 936, July.

Gorton, G. and J. Haubrich, 1986, "Bank Deregulation, Credit Markets and the Control of Capital," Carnegie-Rochester Series on Public Policy (forthcoming).

Hart, P. and B. Holmstrom, 1986, "The Theory of Contracts," working paper #418, M.I.T.

Holmstrom, B., 1979, "Moral Hazard and Observability," Bell Journal of Economics 10, 74-91.

Kane, E., 1985, "The Gathering Crisis in Federal Deposit Insurance," MIT Press, Cambridge.

Litzenberger, R. and J. Van Horne, 1978, "Elimination of the Double Taxation of Dividends and Corporate Financial Policy," Journal of Finance 33, June, 737-757.

Marcus, A., 1983, "The Bank Capital Decision: A Time Series-Cross Section Analysis," Journal of Finance 38, September, 1217-1232.

Melvin, D., 1986, "A Primer for RMA Staff on Legal and Regulatory Concepts and Standards in the Securitization of Loans," Robert Morris Associates, Philadelphia.

Merton, R., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," Journal of Finance 29, May, 449-470.

Miller, M., 1977, "Debt and Taxes," Journal of Finance 30, June, 261-75.

Pavel, C., 1986, "Securitization," Economic Perspectives, Federal Reserve Bank of Chicago.

Pennacchi, G., 1987, "A Reexamination of the Over (or Under) Pricing of Deposit Insurance," Journal of Money, Credit and Banking (forthcoming).

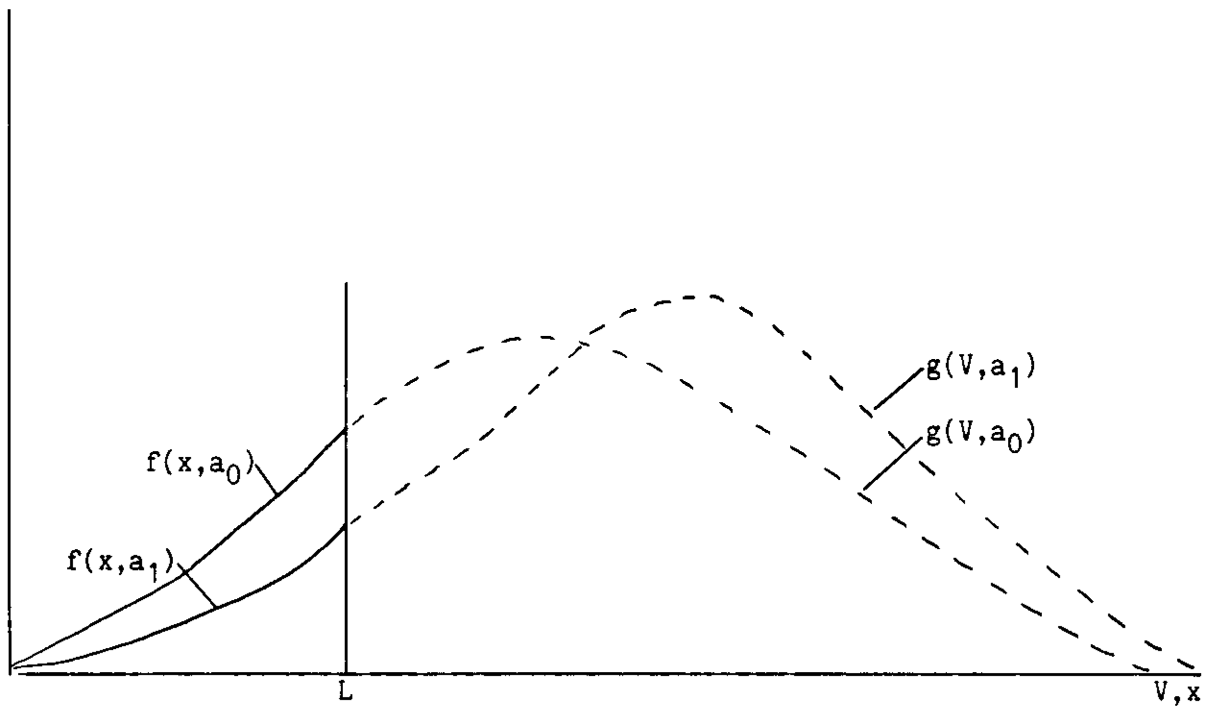
Salem, G., 1985, "Selling Commercial Loans: A Significant New Activity for Money Center Banks," The Journal of Commercial Bank Lending, April, 2-19.

Skelton, J., 1982, "The Relative Pricing of Tax-Exempt and Taxable Debt," Working Paper: University of California-Berkeley.

Trczinka, C., 1982, "The Pricing of Tax-Exempt Bonds and the Miller Hypothesis," Journal of Finance 37, September, 907-924.

Figure 1

Firm Asset (V) and Loan Return (x)
Probability Density Functions



- - - g(V,a) = density of firm's assets

_____ f(x,a) = density of loan return

L = promised loan payment

a_1, a_0 = monitoring levels, where $a_1 > a_0$