

The definitive version is available at wileyonlinelibrary.com.

The Optimal Capital Structure of a Liquidity-insuring Bank

by

HANS-JÜRIG BÜTTLER
Research Group
Swiss National Bank #
and
Department of Economics
University of Zurich
Switzerland

ABSTRACT

This paper deals with the question of the optimal capital structure for a banking firm. It considers a competitive bank as an insurer of unpredictable liquidity demanded by depositors in the sense described by Diamond and Dybvig. The model developed in this paper is able to explain the stylized facts. It considers three key features of a bank: first, the demand deposit contract allows the depositor to run the bank if he believes that the bank's solvency is insufficient. Secondly, four financial states of the shareholders' wealth are considered explicitly. Thirdly, the maximization of the shareholders' expected utility of pay-offs is constrained by the (weakest) condition that the expected yield on equity exceeds the risk-free rate of interest (the 'yield-on-equity constraint'). In the framework of our model, the unconstrained optima as considered in the existing banking literature imply, first, a high equity-to-debt ratio in the order of magnitude of 0.88 to 1 as well as a loan-to-cash ratio of one. Secondly, the bank is not exposed to any risk at all: it stays both solvent and liquid. When the 'yield-on-equity constraint' is taken into account, the constrained optimum may not be contained in the set of unconstrained optima if, for instance, the interest rate differential is small or if the volatility of changes in deposits is large. Such a constrained optimum implies, first, a low equity-to-debt ratio in the order of magnitude of 0.04 as well as a loan-to-cash ratio of one. Secondly, the bank is now exposed to the risk of an insolvency or an illiquidity.

Appeared in *The Econometrics Journal*, Vol. 2, No. 2, 1999, 268 – 291.

Mailing address: Swiss National Bank, 8022 Zurich, Switzerland. Phone (direct dialling): +41-1-631 34 17, Fax: +41-1-631 39 01.

1 Introduction

After the Modigliani and Miller [1958] proposition of the irrelevancy of the capital structure and the vast literature that has evolved thereafter in more than three decades, the task to explain the optimal capital structure of a firm still remains a challenge. Both from a theoretical and regulatory point of view, it is important to know which factors determine the capital structure of a banking firm.

The proposition of the irrelevancy of a firm's capital structure seems to be immune to violations within the perfect market paradigm. Those factors discussed in the literature which are able to overcome the irrelevancy theorem can be grouped in *tax*-related factors, *bankruptcy costs*-related ones and in factors related to *information costs* (agency problem, signalling, moral hazard etc.). The literature, however, has relied too frankly on the trade-off between the tax-shield (which favours debt financing) and expected bankruptcy costs (which favours equity financing and hence limits the advantage of debt). The problem with taxes is that taxes may well influence the optimal capital structure in the aggregate but still leave the individual firm indifferent to its leverage (Miller [1977] or Modigliani [1988]). Although bankruptcy is a widely used argument against the irrelevancy theorem, it is not a very convincing one in the light of the argument of Haugen and Senbet [1978], who show that expected bankruptcy costs are limited to the cost of a financial restructuring that prevents liquidation. These costs are small since they do not include liquidation costs which are normally the biggest part of bankruptcy costs. On the other hand, Haugen and Senbet conclude that the decision about liquidation should be made independently of the financial state of the firm. Rather, the firm should be liquidated only on condition that its liquidation value is higher than its going concern value. Although the likelihood of financial distress does increase as leverage increases, costs of liquidation are not part of the expected costs of financial distress, and therefore should not be considered in the decision about the capital structure. Haugen and Senbet's reasoning is consistent with Warner's [1977] empirical findings of low bankruptcy costs for railway companies.

With respect to the *banking* firm, the researcher finds himself in a even more uncomfortable situation, because not only the liability side of the balance sheet is subject to the irrelevancy theorem but the asset side as well. Assets of a bank can be thought of as a financial portfolio that can be replicated by any investor. Hence, the search for an optimal capital structure for a bank should give a reason for the existence of a bank at the outset as was noted by Santomero [1984, S. 593]. The relevant banking literature mainly relies on bankruptcy costs without clearing up the compelling argument of Haugen and Senbet (see, e. g., Fama [1980], Talmor [1980], Baltensperger [1980]). Furthermore, taxes have been used to derive an optimal capital ratio as in Orgler and Taggart [1983], or market imperfections as in Pringle [1974]. The problem with the market-imperfection approach is that banks would disappear if, and when, these

market imperfections disappear.¹ Some of the literature treats deposits as a bank-specific product and bases the optimal capital structure on an implicit cost function for deposits (see, e. g., Fama [1985], Orgler and Taggart [1983], Baltensperger and Milde [1987]). These explanations of the optimal capital structure, first, suffer from the specific assumptions on the curvature of the cost function and, secondly, lack a discussion about the economic value of deposits. Some of the literature is focused on a particular issue of the banking firm such as bank runs (see, e. g., Jacklin and Bhattacharya [1988], Chari and Jagannathan [1988], or Gorton [1985]). The rest of the capital-related literature is, to our knowledge, primarily concerned with the question of regulation of equity.

The purpose of this paper is to explain the optimal capital structure for a *competitive* banking firm.² It considers the bank as an insurer of unpredictable liquidity demanded by depositors in the sense described by Diamond and Dybvig [1983]. To motivate the design of our model presented in the next section, let us explain the nature of the *demand deposit contract* in more details. As in Diamond and Dybvig [1987], banks sell liquidity insurance against unpredictable consumption demand of depositors more efficiently than a competitive claim market.³ We feel that there are two reasons why banks are more efficient than this market for claims. Since a demand deposit can be interpreted as an American-type puttable bond with floating-rate coupon and with a strike price equal to the face value, the pricing of these puttable bonds is rather difficult.⁴ Moreover, transaction costs would be considerable.

Obviously, the economic value of a demand deposit is equal to the sum of the interest on its holdings plus a risk premium for the insurance the depositor acquires against unpredictable liquidity demand. This premium is paid implicitly by the depositor for the opportunity to withdraw any amount up to his holdings at *face value any time*.⁵ In macroeconomic terms the provision of liquidity insurance finds its equivalent expression in the difference between the monetary base and the next higher aggregate M_1 . The bank's costs to issue deposits include, besides interest payments, real factor costs such as wages, rents for buildings and equipment, safety devices, electronic data processing equipment etc.

The use of the deposit contract has the following four implications. First, if we view deposits as an instrument with which services are sold, deposits must not be considered as an *output* only, but as an *input* as well like the assets of a non-banking firm. To recognize bank liabilities as a factor of production clearly violates the most crucial assumption of Modigliani and

¹ However, it would be compelling to explain why such inefficiencies, if they gave rise to the existence of banks, have prevailed for hundreds of years.

² We do not think that the existence of a banking firm can be explained entirely by non-competitiveness.

³ For a discussion of the "sequential service constraint" in the Diamond and Dybvig [1983] model see for example russel [1993] and McCulloch and Yu [1993]. Further theoretical support for the use of demandable debt is provided by Calomiris and Kahn [1991] and Lacker [1991].

⁴ Since there is no closed-form solution of the American-type puttable bond available, the pricing has to be done numerically (see e. g. Büttler and Waldvogel [1996]).

⁵ We do not negate other deposit-related services such as free (or priced below costs) payment transactions, accounting services and so forth.

Miller [1958], namely that investment decisions are made independently of financing decisions. Hence, we propose a new *multi-dimensional cost function* in our model which treats deposits, loans and cash balances as an output, while input prices (wages, rental price of capital etc.) are assumed to be fixed. It underlies non-increasing economies of scale, the latter being in accordance with a competitive firm.⁶

Secondly, the deposit contract promises a payment on demand at par as long as *confidence* is maintained. Confidence is based primarily on the solvency of the bank. If solvency becomes questionable, the payment to a depositor becomes contingent on his position in line (sequential service constraint), and if he acts rationally, he should run the bank. This, in turn, forces the bank into illiquidity. Hence, *solvency and liquidity are interrelated*. The stochastic nature of deposits does not seem to be a danger to the soundness of a bank in the first place. In a run, when confidence is lost, deposit withdrawals are no longer random: all the deposits are withdrawn, if depositors are fully informed about the status of solvency. In other words, the bank may be liquidated by depositors due to their loss of confidence, rather than by the owners who base their decision on the ratio of the going concern value to the liquidation value of assets. In our view, the expected cost of bankruptcy increase considerably with a higher leverage, in contrast to Haugen and Senbet [1978].⁷ To account for this phenomenon in our model, we consider both a *depositor's reaction function*, which relates solvency with liquidity of the banking firm, and four different *financial states* of the shareholders' end-of-period wealth, which emerge from the possible combinations of solvency/insolvency and liquidity/illiquidity of the banking firm under consideration.

Thirdly, the level of *information* as regards the solvency of a bank is, therefore, crucial in explaining the depositor's behaviour. For this reason, we will introduce a parameter in our model, denoted by α , which measures the *degree of 'informativeness'* of the depositors.

Fourthly, the management of a bank can influence the *probability of a run* by its decision on the capital ratio, the riskiness of assets and cash balances. Hence, the capital structure of a bank has an impact on the probability of both a financial crisis and a liquidation. In our model, the *joint probability density function for deposits and loans*, both prevailing at the end of the period under consideration, depends on the bank manager's decision on these two items at the beginning of a model period.

In summary, our model considers three key features: first, the demand deposit contract allows the depositors to run the bank if they believe that the bank's solvency is insufficient. Secondly, four states of the owners' end-of-period wealth are explicitly considered, namely those states of wealth which emerge from the possible combinations of solvency/insolvency and liquidity/illiquidity of a banking firm. These two key features are modelled by means of the

⁶ The cost function we propose exhibits both increasing and decreasing economies of scale. For the purpose of our investigation, the range has been restricted to non-increasing economies of scale.

⁷ Higher bankruptcy costs for banks is in line with empirical findings of James [1991] who tested for bankruptcy costs of failed US banks.

following building blocks: a depositors' reaction function, the depositors' degree of informativeness, four different financial states of the owners' end-of-period wealth, a joint probability density function for terminal deposits and terminal loans, and a multi-dimensional cost function with non-increasing economies of scale. Thirdly, the optimization is subject to the (possibly weakest) condition that the expected yield on equity exceeds the risk-free rate of interest. We call this the 'yield-on-equity constraint' which was not yet considered in the existing literature.

The stochastic framework has been set and, in principle, opens up the way for practitioners to use it. The model is more realistic than the frictionless (no-constraints) model of Diamond and Dybvig which it generalizes.

The paper is organized as follows. In the second section, we develop the model. The third section presents the results of the numerical sensitivity analysis, followed by conclusions.

2 The Model

The model considered in this paper is, in principle, a one-period model for a negligibly small bank in a competitive market. Variables prevailing at the beginning of the period considered for the model, henceforth the 'model period', do not carry a superscript, while those prevailing at the end of the model period carry an asterisk (*); see the appendix A for a list of variables, functions and parameters. As an exception to be explained later, deposits are treated as if there were two periods. At the beginning, the bank's assets consist of cash balances (C) — which may include reserves held with the central bank — and loans (L). The liabilities considered in this paper consist of deposits (D) and equity (E). Thus, the balance sheet equation reads $C + L = D + E$ at the beginning of the model period. To be more specific, L denotes the going concern value of loans and E the market value of equity.

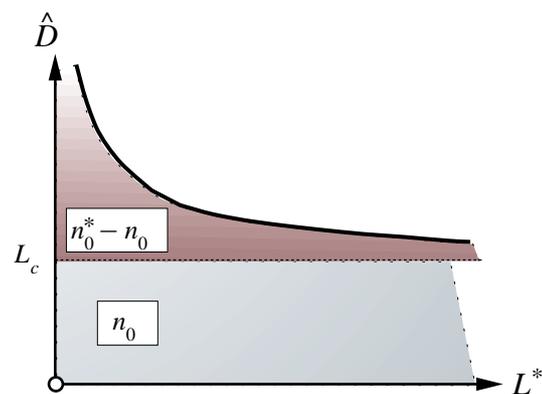
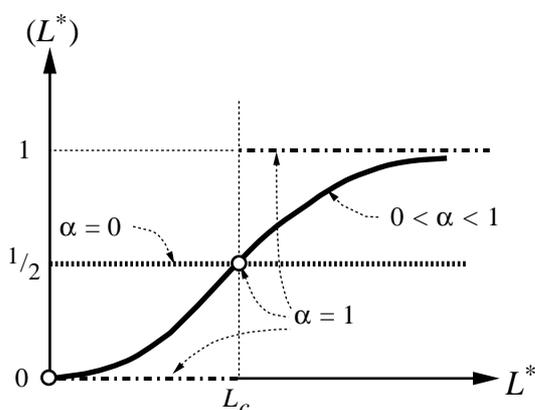


Figure 1: Reaction Function of Depositors **Figure 2:** Ex-post Probability of Illiquidity

During the model period, loans yield a return of r_ℓ and deposits yield an explicit interest rate of r_d , both given exogenously. The provision of liquidity insurance causes (real) factor

costs to the bank in the amount of $\varphi(L, D, C)$. We assume that the marginal cost of producing one unit of loans is higher than that for deposits which, in turn, is higher than that for cash balances for equal loans, deposits and cash balances, that is, $\partial\varphi/\partial L > \partial\varphi/\partial D > \partial\varphi/\partial C > 0$ for $L = D = C$. The evolution of both deposits and loans during the model period is not under the control of the manager-owner of the bank: deposits may be withdrawn by depositors at random, while the going concern value of loans may be re-assessed by the market. We assume, therefore, that deposits and loans, which prevail at the end of the model period, have a joint probability density function, $f(\hat{D}, L^*)$, where \hat{D} denotes the *intermediate* deposit balances before depositors react to the observed financial status of the bank in a second round. Since we wish to focus on the optimal capital structure for a fixed total sum of assets (or liabilities) which is given at the beginning of the model period, we do not consider a possible dependency of this probability density function on loan or deposit activities of the bank. Rather, we assume that this density function is a datum for the bank. It should have two properties: first, neither deposits nor loans may become negative, and secondly, the outcome at the end of the model period $\{\hat{D}, L^*\}$ should depend on the values which the management chooses at the beginning of the model period $\{D, L\}$. A possible candidate is the joint log-normal probability density function which we will consider in the numerical simulations.

In a second but infinitesimally small period of time, depositors are allowed to withdraw all or part of their deposits after they have assessed the status of solvency of the bank under consideration. We assume that these withdrawals of deposits have solely an impact on cash balances. Hence, neither the status of solvency nor any item of the profit and loss account change after depositors have taken an action. This is a strong assumption as soon as we abandon the assumption that the bank under consideration is negligably small, because the withdrawals of deposits force the bank to liquidate its assets quickly which, in turn, induces the selling price, most likely, to be lower than the going concern value. Since the true status of solvency of the bank may not be perceived precisely, we assume that depositors withdraw their deposits according to the following reaction function:⁸

$$D^* = \hat{D} \mathfrak{h}(L^*), \quad \text{with } \mathfrak{h}(L^*) \equiv \left[1 - \frac{1}{1 + (L^*/L_c)^{\text{arctanh}(\alpha)}} \right], \quad \text{where } 0 \leq \alpha \leq 1. \quad (2.1)$$

The *final* amount of deposits (D^*) is a fraction ($0 \leq \mathfrak{h}(L^*) \leq 1$) of the intermediate deposits (\hat{D}) which have been realized before the depositors take an action (see Figure 1). The parameter α ($0 \leq \alpha \leq 1$) serves as a measure of the depositors' information level regarding the solvency of the bank (arctanh denotes the inverse hyperbolic tangent function). The regular case ($0 < \alpha < 1$) is shown as the solid 'log-logistic' curve in the figure 1. From the equation (2.1) emerge two polar cases for $\alpha \rightarrow 1$ and $\alpha \rightarrow 0$. If depositors are fully informed α takes on the value of one;

⁸ Büttler and Schiltknecht [1983] use a similar approach in investigating monetary policy impacts on bank behaviour: in a one-period model, the probability density functions of loans and deposits changes in an implicit second round whenever the monetary policy changed.

this case is shown as the two dashed-dotted horizontal lines including the single point with coordinates $(L_c, 1/2)$ in the figure 1. It tells you that the final deposits are equal to the intermediate deposits if the assessed loans are higher than a critical loan level (L_c) to be explained later, or the bank's depositors withdraw all their deposits if loans are less than this critical loan level, or equal to one-half the intermediate deposits if loans are (exactly) equal to the critical loan level just mentioned. In the latter case, one-half of the bank's depositors keep all their intermediate deposits, while the other half withdraws all of its deposits. In the other polar case when $\alpha = 0$, depositors have no information at all. This case is depicted as the horizontal dotted line in the figure 1. For all loan levels, one-half of the bank's depositors keep all their intermediate deposits, while the other half withdraws all of its deposits.

Since the *intermediate* deposits and the final loans have the joint probability density function $f(\hat{D}, L^*)$, so have the *final* deposits and loans a joint probability density function, $f^*(D^*, L^*)$, in terms of the original density function $f(\cdot, \cdot)$:

$$f^*(D^*, L^*) = \frac{1}{h(L^*)} f\left(\frac{D^*}{h(L^*)}, L^*\right) \quad \text{if } 0 < \alpha < 1. \quad (2.2a)$$

The joint probability density function for the case of completely uninformed depositors is given in the equation (2.2b) below

$$f^*(D^*, L^*) = 2 f(2D^*, L^*) \quad \text{if } \alpha = 0, \quad (2.2b)$$

while that for the case of fully informed depositors is given in the next equation.

$$\begin{aligned} f^*(D^*, L^*) &= \delta(D^*) [1 - \mathcal{H}(L^* - L_c)] \int_{\hat{D}=0}^{\hat{D}=\infty} f(\hat{D}, L^*) d\hat{D} \\ &\quad - \delta(D^*) \delta(L^* - L_c) \int_{\ell^*=0}^{\ell^*=L^*} \int_{\hat{D}=0}^{\hat{D}=\infty} f(\hat{D}, \ell^*) d\hat{D} d\ell^* \\ &\quad + \delta(D^*) \delta(L^* - L_c) \int_{\ell^*=0}^{\ell^*=L_c} \int_{\hat{D}=0}^{\hat{D}=\infty} f(\hat{D}, \ell^*) d\hat{D} d\ell^* \\ &\quad + f(D^*, L^*) \mathcal{H}(L^* - L_c) + \delta(L^* - L_c) \int_{\ell^*=L_c}^{\ell^*=L^*} f(D^*, \ell^*) d\ell^* \quad \text{if } \alpha = 1. \end{aligned} \quad (2.2c)$$

Here, $\delta(\cdot)$ denotes the Dirac delta function and $\mathcal{H}(\cdot)$ the Heaviside unit step function. The joint probability density functions (2.2a) – (2.2c) are derived in the appendix B.

The accounting status of the bank is shown in the table 1. In the profit and loss account for the model period, cash flows originate from interest payments and real factor costs. The change in the value of loans (ΔL) is assumed to have no impact on cash balances. In the balance sheet prevailing at the end of the model period, both cash flows and changes in deposits during the model period are reflected in the change of cash balances. Hence, we assume that changes in deposits do not have an impact on loans. Recall, however, that final (or intermediate) deposits

and loans prevailing at the end of the model period have a joint probability density function ($f(\hat{D}, L^*)$ or $f^*(D^*, L^*)$, respectively). The change in equity reflects profits (P) or losses ($-P$) during the model period.

Table 1: Accounting Status for the ‘Model Period’

Balance Sheet at the End of the Model Period		Profit and Loss Account for the Model Period	
Assets	Liabilities	Expenditures	Earnings
$C^* = C + r_\ell L - r_d D$	$D^* = D + \Delta D$	$r_d D$	$r_\ell L$
$- \varphi(L, D, C) + \Delta D$		$\varphi(L, D, C)$	ΔL
$L^* = L + \Delta L$	$E^* = E + \Delta E = E + P$	P	
$C^* + L^*$	$D^* + E^*$	$r_d D + \varphi(L, D, C) + P$	$r_\ell L + \Delta L$

We assume that the manager is maximizing the owner’s utility of final wealth. The terminal value of equity becomes from the accounting identity:

$$E^* = C^* + L^* - D^* = L^* - L_c, \quad \text{with } L_c \equiv (1+r_d)D + \varphi(L, D, C) - r_\ell L - C. \quad (2.3)$$

By L_c we denote the critical level of the value of loans the bank needs at least to keep its solvency. After the reaction of depositors on which the management has no influence, the bank is left with one of the four financial states described in the table 2.⁹ The bank is insolvent if the terminal value of equity is negative, while it is illiquid if the terminal cash balances are negative. Denote for a moment all the intermediate variables with a ‘hat’ ($\hat{\cdot}$), then the bank is insolvent if $\hat{E} = \hat{L} + \hat{C} - \hat{D} < 0$. Given the depositors’ reaction in the second round as described before, this implies $E^* = L^* + C^* - D^* < 0$ which, in turn, means that $L^* < L_c$ (as shown in the equation (2.3) above). The bank is illiquid if $C^* < 0$ which implies $D^* < L_c$. *In summary, insolvency implies $L^* < L_c$ and illiquidity implies $D^* < L_c$.*

Table 2: Pay-offs in the Four Financial States Considered

	solvent and ...		insolvent and ...	
	... liquid (s=1)	... illiquid (s=2)	... liquid (s=3)	... illiquid (s=4)
$[A_s]_{\alpha=1}$	E^*	$\max [E^* - S, 0]$ ^(a)	0 ^(c)	0 ^(c)
$[A_s]_{\alpha=0}$	E^*	0 ^(b)	$\max [r_\ell L - r_d D - \varphi(L, D, C), 0]$ ^(d)	0 ^(c)
A_s	E^*	$\max [\alpha(E^* - S), 0]$	$\max [(1 - \alpha) \{r_\ell L - r_d D - \varphi(L, D, C)\}, 0]$	0

Notes for Table 2 :

- ^a Liquidity can be restored in the interbank market due to solvency. S denotes penalty cost.
- ^b The banking firm goes bankrupt because the going concern value of loans is assumed to be higher than the value in liquidation.
- ^c The banking firm is assumed to become liquidated.
- ^d The manager-owner pays dividends out of cash flows.

⁹ Note that each of the four states described in the Table 1 could be complemented with two additional states which distinguish between the liquidation and the continuation of the bank. Such a distinction, however, makes only sense if the value in liquidation can be determined endogeneously, a task which is beyond our paper. We neglect, therefore, these additional states of the bank.

In the table 2, the pay-off A_s for one of the four financial states ($s = 1, \dots, 4$) which will be received by the owners at the end of the model period is assumed to be a linear combination of those two pay-offs which prevail if the depositors are either fully informed ($\alpha = 1$) or completely uninformed ($\alpha = 0$):

$$A_s = \alpha [A_s]_{\alpha=1} + (1-\alpha) [A_s]_{\alpha=0} \quad \text{for } s = 1, \dots, 4. \quad (2.4)$$

Observe that only the pay-offs in the first two states are random variables. They depend on the random value of equity prevailing at the end of the model period which, in turn, depends on the random value of loans prevailing at the end of the model period, that is, $A_s = A_s(E^*) = A_s(E^*(L^*))$. Let us comment briefly on the four financial states considered in the table 2. First, if the bank is both *solvent and liquid* at the end of the model period, then the pay-off is equal to the value of the equity in any case ($s = 1$ and any α). Secondly, the *illiquid but solvent* bank can restore its liquidity in the interbank market (or with the help of the central bank in question) at a higher interest rate than the prevailing market rates conditional on the fact that depositors know (with certainty) that their bank is solvent (see the case (a) in the table 2). Hence, the bank faces penalty cost S assumed to be fixed in our model. If depositors are completely uninformed, then the pay-off depends on the value of the loans in liquidation, denoted by L_ℓ^* :

$$[A_2]_{\alpha=0} = \begin{cases} E^* & \text{if } L_\ell^* > D^*, \\ 0 & \text{if } L_\ell^* \leq D^*. \end{cases} \quad (2.5)$$

We assume that the going concern value of loans is considerably higher than the value in liquidation, hence $L_\ell^* \leq D^*$ or $[A_2]_{\alpha=0} = 0$ (see the case (b) in the table 2). Another argument to support this assumption is that illiquidity is, from the point of view of the non-informed depositor, a sign of insolvency which may lead him to run the bank. In this case all loans need be liquidated at a great pace. Thirdly, if the bank is *insolvent irrespective of its status of liquidity*, the pay-off is, in principle, zero (see the three cases (c) in the table 2). Fourthly, if the bank is *insolvent but liquid*, it is conceivable, however, that the manager-owner has the opportunity to redistribute wealth from depositors to himself by paying dividends out of cash flows if the bank's depositors are completely uninformed about the financial status (see the case (d) in the table 2). Due to the reaction function $h(\cdot)$ the probability of the third state is very small if depositors are completely informed ($s = 3$ and $\alpha = 1$).

The goal of the risk-averse manager-owner is to maximize the expected utility received from the end-of-period value of his residual claim on the value of the firm given by the pay-offs A_s in the four states:

$$\begin{aligned} \max_{\{C, L, D, E\}} \mathcal{E} \mathcal{U}(A) &= \max_{\{C, L, D, E\}} \sum_{s=1}^4 \mathcal{E}_s \mathcal{U}(A_s(L^*)), \\ \text{s. t. (1) } C, L, D, E &\geq 0, \quad (2) \mathcal{E} \left\{ \frac{A_s(E^*)}{E} - 1 \right\} \geq r_f > r_a, \quad (3) C + L = D + E = 1, \quad (4) L_c \geq 0, \end{aligned} \quad (2.6)$$

with the meaning that \mathcal{E} denotes the expectation operator and \mathcal{U} the concave utility function. When we deal with risk-neutral owners, the marginal utility (of the then linear utility function) is set equal to one, that is, $\mathcal{U}'(\cdot) = 1$ without loss of generality. The maximization of the expected utility of the four pay-offs considered in our model is subject to four constraints given in the equation (2.6) above. First, all the four instruments chosen by the manager-owner, i. e., the cash balances, the loan level, the deposit accounts, and the equity proportion should be non-negative. Secondly, an investor can become either a depositor or an owner of the bank by assumption (but not both). Following Arrow [1971], the expected rate of return on equity should exceed the risk-free interest rate, denoted as r_f , if the owner of the bank is risk averse. Furthermore, since the demand deposit contract provides a risk-averse depositor with a liquidity insurance, the risk-free interest rate should exceed the rate of interest on sight deposits by the amount of the risk premium at least. In summary, this gives rise to the second constraint $\mathcal{E}\{A_s(E^*)/E - 1\} \geq r_f > r_d$, henceforth the ‘yield-on-equity constraint’. In our view, this is the weakest condition one can impose on the yield on equity.¹⁰ There is yet another way of looking at the ‘yield-on-equity constraint’. In a principal-agent problem, the principal may not know much about the probabilities of solvency because of the information costs, but is able to observe the ‘yield-on-equity constraint’, and will base his decisions mainly on such ratios. This is not far in spirit from the ‘degree of informativeness’ which we introduced in the model by the parameter α .¹¹ Thirdly, the balance sheet identity should hold. Since we are interested in the substitution effects between both assets or between both liabilities, we restrict the total balances of our bank to be one unit. Fourthly, the manager of the bank could, in principle, choose a negative critical loan level. However, this does not improve the expected pay-offs because the random values of both deposits and loans cannot become negative at the end of the model period.

The manager’s objective function (2.6) can be written more explicitly with the help of the joint probability density function for the final deposits and loans. Recalling that the terminal value of loans and the terminal value of deposits each must be greater than the critical loan level in order that the bank is both solvent and liquid, the expected utility of the terminal pay-offs can be written as

¹⁰ Two more elaborated extensions of the ‘yield-on-equity constraint’ are conceivable at least. In the first one, the shareholder of the bank considered is also allowed to hold public debt as well as common shares of non-bank firms. The expected yield on equity of the banking firm in question which would arise in a general equilibrium framework should be the lower limit for the optimization given in the equation (2.6). This general equilibrium approach, however, is beyond the scope of this paper. In a second extension, the expected rate of return on the bank’s assets is the weighted sum of the expected rates of return on the two liabilities of the bank considered, the weights being the ratios of the respective liabilities to the total balance sheet. If one assumes that the Capital Asset Pricing Model (CAPM) holds, the expected rate of return on the bank’s assets can be substituted by the well-known CAPM relationship in terms of the beta-risk measure and the expected rate of return on the market portfolio. Taking these two relationships together, the expected rate of return on equity in the capital asset pricing framework becomes a function of the beta-risk measure, the expected rate of return on the market portfolio, the risk-free interest rate and the debt-to-equity ratio. This ‘CAPM yield-on-equity constraint’, however, does not change the pattern of the optima as explained in the third section of this paper.

¹¹ I am indebted to Karim Abadir for pointing this out to me.

$$\begin{aligned}
\mathcal{E} \mathcal{U}(A) &= \int_{L^*=L_c}^{L^*=\infty} \int_{D^*=L_c}^{D^*=\infty} \mathcal{U}(A_1) f^*(D^*, L^*) dD^* dL^* && \text{(state 1)} \\
&+ \int_{L^*=L_c}^{L^*=\infty} \int_{D^*=0}^{D^*=L_c} \mathcal{U}(A_2) f^*(D^*, L^*) dD^* dL^* && \text{(state 2)} \\
&+ \mathcal{U}(A_3) \int_{L^*=0}^{L^*=L_c} \int_{D^*=L_c}^{D^*=\infty} f^*(D^*, L^*) dD^* dL^* && \text{(state 3)} \\
&+ \mathcal{U}(A_4=0) \int_{L^*=0}^{L^*=L_c} \int_{D^*=0}^{D^*=L_c} f^*(D^*, L^*) dD^* dL^* && \text{(state 4)}.
\end{aligned} \tag{2.7}$$

Although we may use a transformed version of the equation (2.7) to derive the necessary conditions for an optimal capital structure, the above equation is used in all the numerical simulations to be explained in the next section.

To make clear the role of the depositor's reaction function, note that the probability of a bank insolvency is the same at the intermediate point in time ('ex-ante' probability) as well as at the end of the model period ('ex-post' probability). What changes is the probability of an illiquid status as given below.

$$\begin{aligned}
n_0 &= \int_{\hat{D}=0}^{\hat{D}=L_c} \int_{L^*=0}^{L^*=\infty} f(\hat{D}, L^*) d\hat{D} dL^* \\
n_0^* &= \int_{L^*=0}^{L^*=\infty} \int_{x=0}^{x=L_c/\mathfrak{h}(L^*)} f(x, L^*) dx dL^* > n_0 \quad \text{with } x = \frac{D^*}{\mathfrak{h}(L^*)}
\end{aligned} \tag{2.8}$$

The 'ex-ante' probability of illiquidity (n_0) is shown in the first line and the 'ex-post' probability (n_0^*) in the second line. The redistribution of probability mass from the liquid state to the illiquid state is due to the reaction function $\mathfrak{h}(L^*)$. This relationship is depicted in the figure 2, which shows the corresponding areas of integration.

3 A Numerical Sensitivity Analysis

In the appendix C, we show that the first-order optimality conditions form a non-linear equation system involving two-dimensional integrals. This equation system is not tractable analytically. Although our model is simple, only a numerical sensitivity analysis can unveil the optimal equity-to-debt ratio. In the sequel, we describe the various functions to be used for the numerical sensitivity analysis, and then we present the results along with an intuitive explanation of the optimal equity-to-debt ratio. The basic set of parameter values is listed in the appendix A. The numerical procedure to evaluate the two-dimensional integrals of the equation (2.7) is described in the appendix D.

Consider the joint probability density function of the intermediate deposits and the final loans to be the bivariate log-normal density function:

$$f(\hat{D}, L^*) = \frac{1}{\hat{D} L^* \sigma_d \sigma_\ell} \mathfrak{g} \left(\frac{\ln(\hat{D}) - [\ln(D) + \mu_d]}{\sigma_d}, \frac{\ln(L^*) - [\ln(L) + \mu_\ell]}{\sigma_\ell}, \rho \right), \text{ where} \quad (3.1)$$

$$\mathfrak{g}(x, y, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2} \frac{x^2 - 2\rho xy + y^2}{1-\rho^2}\right).$$

The coefficient of correlation is assumed to range between zero and one ($0 \leq \rho < 1$). Both intermediate deposits and terminal loans (\hat{D}, L^*) are non-negative. Observe that the mean values depend on the deposits and loans (D, L) which have been chosen by the manager at the beginning of the model period. See also the comments for μ_d and μ_ℓ in the appendix A.

Consider the following multi-dimensional cost function (φ) for the n outputs x_1, x_2, \dots, x_n , given fixed input prices which are thus deleted:

$$\varphi(\mathbf{x}) = \delta + \kappa \operatorname{arctanh}(\xi), \quad \text{where } \xi \equiv \theta_1 \|\mathbf{x}\| + \theta_2, \quad \mathbf{x} = (x_1, x_2, \dots, x_n), \quad (3.2)$$

$$\|\mathbf{x}\| = [\beta_1 x_1^m + \beta_2 x_2^m + \dots + \beta_n x_n^m]^{1/m}, \quad 0 \leq a \leq x_i \leq b < \infty, \quad 0 < \beta_i \ (\forall i); \quad 0 < m.$$

This cost function is a linear transformation of the inverse hyperbolic tangent function ($\operatorname{arctanh}$) which is concave in the domain $-1 \leq \xi \leq 0$, but convex in the domain $0 \leq \xi \leq +1$. The parameter δ which shifts the $\operatorname{arctanh}$ function vertically up or down is used to set the cost function equal to zero for a zero output, while the parameter κ stretches the $\operatorname{arctanh}$ function. By assumption, each output (x_i) lies in the finite and positive range $[a, b]$ which is fixed. The one-dimensional output, ξ , is a linear transformation of a weighted-sum norm of the n -dimensional output \mathbf{x} . The exponent m may be any positive real number; the value of one is the borderline between concave and convex iso-cost curves. Because we wish to avoid unbounded costs, the two parameters θ_1 and θ_2 are set in such a way that the output range $[a, b]$ maps, in general, into the truncated domain $[-(1-\varepsilon), 1-\varepsilon]$ of the $\operatorname{arctanh}$ function, where ε may be any positive real number less than one; it may be suitable for most problems at hand that ε is in the order of magnitude of 0.01. The positive parameters β_i determine the magnitude of the marginal costs because

$$\frac{\partial \varphi(\mathbf{x})}{\partial x_i} = \frac{\kappa \theta_1 \beta_i x_i^{m-1} \|\mathbf{x}\|^{1-m}}{1 - [\theta_1 \|\mathbf{x}\| + \theta_2]^2}. \quad (3.3)$$

Notice that the marginal costs depend on *all* the outputs under consideration. Let the first output be the value of loans, the second output be the value of deposits, and the third output be the cash balances, that is, $x_1 = L, x_2 = D, x_3 = C$. Since we assume that the marginal cost of producing one unit of loans is higher than the one for deposits which, in turn, is higher than that for cash balances, that is, $\partial \varphi / \partial L > \partial \varphi / \partial D > \partial \varphi / \partial C > 0$ for $L = D = C$, we must have $\beta_1 > \beta_2 > \beta_3$.

Consider now the cost function of the competitive banking firm explicitly. By assumption, the three outputs (loans L , deposits D , and cash balances C) lie between zero and one, that is, $a = 0$ and $b = 1$. Therefore, the cost function (3.2) may be written as:

$$\varphi(L, D, C) = \delta + \kappa \operatorname{arctanh}\left(\theta_1 [\beta_1 L^m + \beta_2 D^m + \beta_3 C^m]^{1/m} + \theta_2\right). \quad (3.4a)$$

In order to obtain a convex cost function on its whole domain, we map the range of $[a, b]$ into $[0, 1 - \varepsilon]$, that is, we set $0 = \theta_1 [\beta_1 + \beta_2 + \beta_3]^{1/m} a + \theta_2$ and $1 - \varepsilon = \theta_1 [\beta_1 + \beta_2 + \beta_3]^{1/m} b + \theta_2$, which implies

$$\theta_1 = \frac{1 - \varepsilon}{[\beta_1 + \beta_2 + \beta_3]^{1/m}}, \quad \theta_2 = 0, \quad \delta = 0. \quad (3.4b)$$

Observe that δ has been obtained from the assumption that $\varphi(0, 0, 0) = 0$. With these specifications, the cost function (3.4) underlies decreasing economies of scale throughout its domain. An illustration is given for the output vector $\mathbf{x} = (L, D, 1 - L)$ in the figure 3.

Consider next the hyperbolic absolute risk-aversion (HARA) utility function (see e. g. Merton [1971])

$$\mathcal{U}(A) = \frac{1 - \gamma}{\gamma} \left[\frac{\lambda A}{1 - \gamma} + \eta \right]^\gamma, \quad \text{with } \gamma \neq 1, \lambda > 0, \frac{\lambda A}{1 - \gamma} + \eta > 0, \eta = 1 \text{ if } \gamma = -\infty. \quad (3.5)$$

Here A denotes the end-of-period value of wealth, or in our application the different pay-offs. To quote from Merton, “this family of HARA utility functions is rich, in the sense that by suitable adjustment of the parameters, one can have a utility function with absolute or relative risk aversion increasing, decreasing, or constant.” Following Arrow [1971, chapter 3], we assume that, first, the absolute risk aversion ($-\mathcal{U}''(A) / \mathcal{U}'(A)$) decreases with increasing wealth and, secondly, the relative risk aversion ($-\mathcal{U}''(A) A / \mathcal{U}'(A)$) increases with increasing wealth, that is,

$$\begin{aligned} \frac{d}{dA} \left[-\frac{\mathcal{U}''(A)}{\mathcal{U}'(A)} \right] &= \frac{-1}{[1 - \gamma] \left[\frac{A}{1 - \gamma} + \frac{\eta}{\lambda} \right]^2} < 0 \quad \text{for } -\infty < \gamma < 1 \\ \frac{d}{dA} \left[-\frac{\mathcal{U}''(A)}{\mathcal{U}'(A)} A \right] &= \frac{\eta/\lambda}{\left[\frac{A}{1 - \gamma} + \frac{\eta}{\lambda} \right]^2} > 0 \quad \text{for } \eta > 0 \text{ and } -\infty \leq \gamma \leq +\infty, \gamma \neq 1. \end{aligned} \quad (3.6)$$

Moreover, Arrow has shown that the relative risk aversion should be in the order of magnitude of one. An illustration of the HARA utility function for the basic set of parameter values is given in the figure 4. The utility function is shown as the monotonically increasing, bold-faced curve, the absolute risk aversion as the monotonically decreasing, solid curve, and the relative risk aversion as the monotonically increasing, dotted curve.

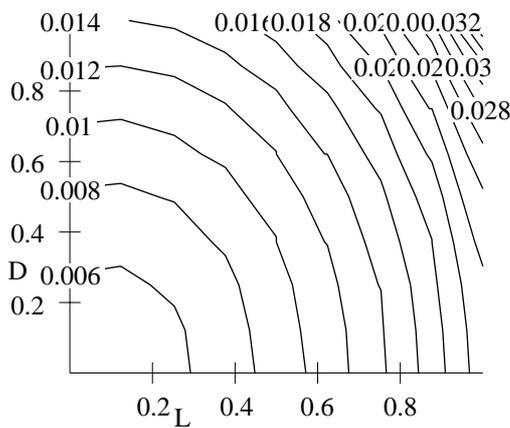


Figure 3: Iso-cost Curves of Function (3.4) for Basic Set of Parameter Values

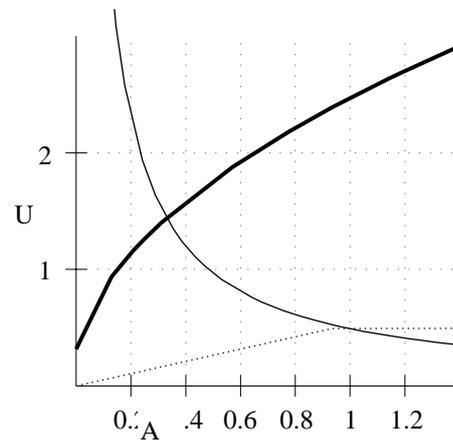


Figure 4: HARA Utility Function (3.5), Absolute and Relative Risk Aversion (3.6)

The numerical sensitivity analysis can be summarized as follows. Assume for a moment that the manager is risk-neutral, that is, he maximizes the expected pay-offs as given in the equation (C.1). Furthermore, assume that this maximization does not underlie the ‘yield-on-equity constraint’. Under these two assumptions just mentioned, the expected pay-offs become the greatest if the manager chooses loans and deposits in such a way that the bank stays both solvent and liquid. The bank stays both solvent and liquid, that is, it enters the first financial state considered in the table 2, if the critical level of loans, as given by the equation (2.3), is less than or equal to zero as shown by the shaded region $OABC0$ in the figure 5a. Within this region, the expected pay-offs are equal to the expected value of the terminal value of loans, that is, $\mathcal{E}\{A_s(E^*)\} = \mathcal{E}\{E^*\} = \mathcal{E}\{L^*\} - L_c = \mathcal{E}\{L^*\}$ by the equation (2.3). Furthermore, the expected value of the terminal value of loans is equal to the value of loans the manager chooses at the beginning of the model period because we have assumed that the balance sheet does not grow, that is, $\mathcal{E}\{L^*\} = L$ (see also the comments for μ_d and μ_ℓ in the appendix A). Therefore, the greatest expected pay-offs in the first financial state occur when the beginning-of-period value of loans is equal to one. Hence, the unconstrained optima occur along the horizontal line between point A and point B in the figure 5a, where the manager chooses to hold no cash and a high equity-to-debt ratio in the order of magnitude of 0.88 to 1. To our knowledge, the ‘yield-on-equity constraint’ is not considered in the existing banking literature; hence, the unconstrained optima as shown in the figure 5a would be the outcome of the known banking models.

It is clear from the monotone transformation of the utility function that the unconstrained optima as described by the equation (2.7) still occur along the horizontal line AB in the figure 5a. It follows that these unconstrained optima cannot explain the low equity-to-debt ratios of banks observed nowadays. Taking the existing banking literature at face value, the manager chooses assets and liabilities in such a way that the bank is *never* exposed to any risk of failure, that is, the bank becomes neither insolvent nor illiquid.

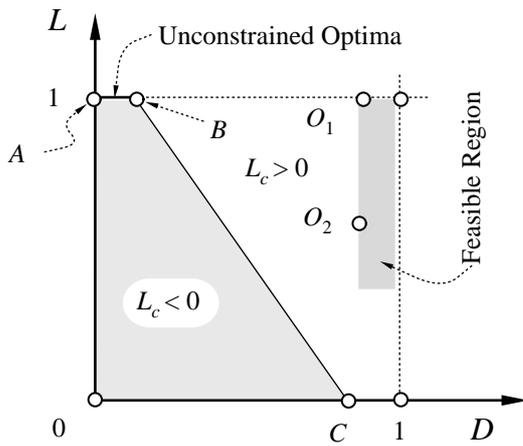


Figure 5a: Constrained Optima for Basic Set of Parameter Values

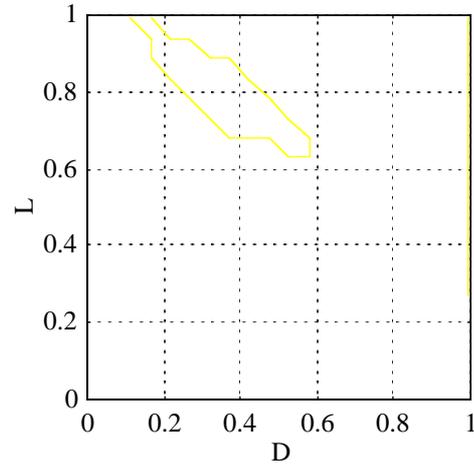


Figure 5b: Two Feasible Regions for Basic Set except $r_\ell = 0.1$

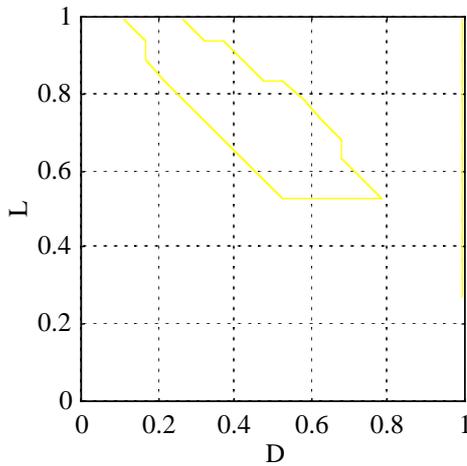


Figure 5c: Two Feasible Regions for Basic Set except $r_\ell = 0.12$

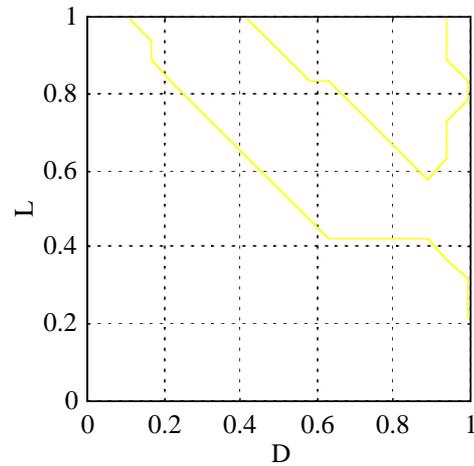


Figure 5d: One Feasible Region for Basic Set except $r_\ell = 0.15$

However, the numerical sensitivity analysis reveals that the unconstrained optima may not be feasible. For the basic set of parameter values considered, the feasible region, as given by the ‘yield-on-equity constraint’ in the equation (2.7), is shown as the shaded rectangle in the figure 5a. It is conceivable that the optimum occurs either at the point O_1 , which prevails in most cases studied for this paper, or at the point O_2 . Hence, the constrained optimum O_1 is in accordance with the empirical facts, namely, a low equity-to-debt ratio and a high loan-to-cash ratio. In fact, the latter ratio is equal to one in most circumstances. Given the fact that cash balances do not bear any interest in the model considered for this paper, this is not surprising at all.

How does the feasible region change as the parameter values change? The figures 5b through 5d demonstrate the modification of the feasible region if the rate of interest on loans increases from 8% to 15%, all the other parameters remaining fixed. Observe that the constrained optima which emerge from the figures 5b through 5d are equal to the unconstrained optimum as shown in the figure 5a, because any point off the region where the critical level of loans is less

than or equal to zero has a lower expected pay-off than any point within this region. This is due to the fact that the pay-off in the first financial state is always greater than in all the other three financial states. We conclude from the numerical analysis that the larger the interest rate differential between assets (loans) and liabilities (deposits), the more likely is an unconstrained optimum as given by the point B in the figure 5a.

Hence, the model is able to explain some stylized facts about the historical evolution of banks. Typically, those banks which have been founded in the nineteenth century started with a high equity-to-debt ratio as well as a high interest-rate differential. In terms of our model, the high equity-to-debt ratio is, from the point of view of depositors, a signal for an unconstrained optimum which, in turn, means that the bank cannot become insolvent nor illiquid (see the figure 5d). By this strategy, the banks were able to build up a reputation. As the competition increased, the interest-rate differential became smaller; hence, the feasible regions fell apart (see the figures 5b & c). Finally, a low equity-to-debt ratio emerges along with a risk to fail (see the constrained optimum O_1 in the figure 5a). From another point of view, the reputation which was built up by the banks in a first historical period may helped them to undergo the exposure to insolvency risk or illiquidity risk.

4 Conclusions

In the framework of our model, the unconstrained optima as considered in the existing banking literature imply, first, a high equity-to-debt ratio in the order of magnitude of 0.88 to 1 as well as a loan-to-cash ratio of one. Secondly, the bank is not exposed to any risk at all: it stays both solvent and liquid.

When the 'yield-on-equity constraint' is taken into account, the constrained optimum may not be contained in the set of unconstrained optima if, for instance, the interest rate differential is small or if the volatility of changes in deposits is large. Such a constrained optimum implies, first, a low equity-to-debt ratio in the order of magnitude of 0.04 as well as a loan-to-cash ratio of one. Secondly, the bank is now exposed to the risk of an insolvency or an illiquidity.

The stochastic framework has been set and, in principle, opens up the way for practitioners to use it. The model is more realistic than the frictionless (no-constraints) model of Diamond and Dybvig which it generalizes.

Acknowledgements

I am indebted to Karim Abadir (co-editor), an anonymous referee, Hans Neukomm, and the participants of the two conferences mentioned below for their valuable comments. This paper has been presented at the 6th Annual Meeting of the European Financial Management Association which was held in Istanbul (Turkey) in June 1997 and at the Conference on

Quantitative Methods in Finance which was held in Sydney, Cairns and Canberra (Australia) in August and September 1997.

Appendix A: List of Parameters, Variables, and Functions

The values given in square brackets are those of the basic set.

Roman Letters

A_s	Pay-offs in the four financial states considered, $s = 1, \dots, 4$.
C	Initial level of cash balances. By assumption, cash balances do not bear any interest.
C^*	Cash balances at the end of the period.
D	Initial level of deposits.
\hat{D}	Intermediate level of deposits before depositors react.
D^*	Amount of deposits at the end of the period after the reaction of depositors.
E	Initial level of market value of equity.
E^*	Market value of equity at the end of the period.
$f(\cdot, \cdot)$	Joint probability density function for intermediate deposits and final loans. A bivariate log-normal density function is used for the simulations.
$f^*(\cdot, \cdot)$	Joint probability density function for final deposits and final loans derived from $f(\cdot, \cdot)$ and the depositors' reaction function.
L	Initial level of the going concern value of loans.
L^*	Value of loans at the end of the period.
L_c	Critical or threshold value of loans. The bank is solvent if its level of loans exceeds this critical value.
L_ℓ^*	Value of loans in liquidation.
m	Degree of norm used for the cost function $\varphi(\cdot, \cdot, \cdot)$. [2].
P	Profit.
r_d	Rate of interest on sight deposits. [0.04].
r_f	Risk-free interest rate. [0.06].
r_ℓ	Rate of interest on loans. [0.08].
S	Penalty cost for illiquid but solvent bank. [0.16].

Greek Letters

α	Level of information of depositors. [0.9].
β_i	Weights of norm used for the cost function $\varphi(\cdot, \cdot, \cdot)$, $i = 1, \dots, n$. The weights determine the marginal costs with respect to each of several outputs. [1, 0.5, 0.1].
$\delta(\cdot)$	DIRAC delta function.
δ	Shift parameter of the cost function $\varphi(\cdot, \cdot, \cdot)$.
ε	Parameter of the cost function $\varphi(\cdot, \cdot, \cdot)$. [0.01].
γ	Parameter of the HARA utility function. [0.5].

λ	Parameter of the HARA utility function. [3].
$\varphi(\cdot, \cdot, \cdot)$	Cost function. An arctanh function with decreasing economies of scale is used for the simulations.
η	Parameter of the HARA utility function. [0.1].
κ	Stretching parameter of the cost function $\varphi(\cdot, \cdot, \cdot)$. [0.02].
μ_d	Mean of the logarithm of the deposit rate, $\ln(\hat{D} / D)$. Given the geometric Brownian motion $d\hat{D} / \hat{D} = \hat{\mu}_d dt + \sigma_d dz$ with time t and Wiener process z , no growth implies that $\hat{\mu}_d = 0$ and $\mu_d \equiv \hat{\mu}_d - \sigma_d^2 / 2 = -\sigma_d^2 / 2$.
μ_ℓ	Mean of the logarithm of the loan rate, $\ln(L^* / L)$. Given the geometric Brownian motion $dL^* / L^* = \hat{\mu}_\ell dt + \sigma_\ell d\zeta$ with time t and Wiener process ζ , no growth implies that $\hat{\mu}_\ell = 0$ and $\mu_\ell \equiv \hat{\mu}_\ell - \sigma_\ell^2 / 2 = -\sigma_\ell^2 / 2$. The two Wiener processes, z and ζ , are correlated with coefficient ρ .
π	Circle ratio. $\pi = 3.14159\dots$
ρ	Coefficient of correlation between rates of return on deposits and rates of return on loans. [0].
σ_d	Variance of rates of return on deposits. [0.3].
σ_ℓ	Variance of rates of return on loans. [0.2].
θ_i	Parameters of the cost function $\varphi(\cdot, \cdot, \cdot)$, $i = 1, 2$.
ξ	Auxiliary Variable of the cost function $\varphi(\cdot, \cdot, \cdot)$.

Special Symbols

\mathcal{E}	Expectation operator.
$\mathbf{g}(\cdot, \cdot, \cdot)$	Auxiliary function for the bivariate log-normal density function.
$\mathbf{h}(\cdot)$	Reaction function of depositors according to assessment of solvency.
$\mathcal{H}(\cdot)$	Heaviside unit step function.
$\ \mathbf{x}\ $	Norm defined as $[\beta_1 x_1^m + \beta_2 x_2^m + \dots + \beta_n x_n^m]^{1/m}$ with $\mathbf{x} = (x_1, x_2, \dots, x_n)$. It is used for the multi-dimensional cost function $\varphi(\cdot, \cdot, \cdot)$.
\mathcal{U}	Concave utility function. A HARA utility function is used for the simulations.

Appendix B: Derivation of the Terminal Joint Probability Density Function

In this appendix, we derive the joint probability density function for final deposits and final loans, $f^*(\cdot, \cdot)$, from the joint probability density function for intermediate deposits and final loans, $f(\cdot, \cdot)$, and the depositors' reaction function, $h(\cdot)$. First, we consider the regular case for the level of informativeness, that is, $0 < \alpha < 1$. Recalling that the final deposits are given by $D^* = \hat{D} h(L^*)$ and recalling that the inverse of $h(\cdot)$ exists, the joint probability function for final deposits and final loans, $F^*(\cdot, \cdot)$, can be written as:

$$\begin{aligned} F^*(D^*, L^*) &= \text{Prob}[d^* \leq D^*, \ell^* \leq L^*] = \text{Prob}\left[\hat{d} \leq D^*/h(\ell^*), \ell^* \leq L^*\right] \\ &= \int_{\ell^*=0}^{\ell^*=L^*} \int_{\hat{d}=0}^{\hat{d}=D^*/h(\ell^*)} f(\hat{d}, \ell^*) d\hat{d} d\ell^*. \end{aligned} \quad (\text{B.1})$$

The joint probability density function for final deposits and final loans, $f^*(\cdot, \cdot)$, is the partial derivative with respect to both random variables, i. e.,

$$f^*(D^*, L^*) = \frac{\partial^2 F^*(D^*, L^*)}{\partial L^* \partial D^*} = \frac{1}{h(L^*)} f\left(\frac{D^*}{h(L^*)}, L^*\right). \quad (\text{B.2})$$

This is the equation (2.2a) in the text. Next, consider the case when the depositor is completely uninformed, that is, $\alpha = 0$. The joint probability function for final deposits and final loans, $F^*(\cdot, \cdot)$, can be written now as

$$\begin{aligned} F^*(D^*, L^*) &= \text{Prob}[d^* \leq D^*, \ell^* \leq L^*] = \text{Prob}[\hat{d} \leq 2D^*, \ell^* \leq L^*] \\ &= \int_{\ell^*=0}^{\ell^*=L^*} \int_{\hat{d}=0}^{\hat{d}=2D^*} f(\hat{d}, \ell^*) d\hat{d} d\ell^*, \end{aligned} \quad (\text{B.3})$$

because $h(\cdot) = 1/2$ for any value of loans. Differentiation of the equation (B.3) with respect to both random variables yields the equation (2.2b) in the text. Finally, consider the case when the depositor is completely informed, that is, $\alpha = 1$. The joint probability function for final deposits and final loans, $F^*(\cdot, \cdot)$, can be written now as

$$\begin{aligned} F^*(D^*, L^*) &= \mathcal{H}(D^* - 0) [1 - \mathcal{H}(L^* - L_c)] \int_{\ell^*=0}^{\ell^*=L^*} \int_{\hat{d}=0}^{\hat{d}=\infty} f(\hat{d}, \ell^*) d\hat{d} d\ell^* \\ &+ \mathcal{H}(D^* - 0) \mathcal{H}(L^* - L_c) \int_{\ell^*=0}^{\ell^*=L_c} \int_{\hat{d}=0}^{\hat{d}=\infty} f(\hat{d}, \ell^*) d\hat{d} d\ell^* \\ &+ \mathcal{H}(L^* - L_c) \int_{\ell^*=L_c}^{\ell^*=L^*} \int_{\hat{d}=0}^{\hat{d}=D^*} f(\hat{d}, \ell^*) d\hat{d} d\ell^*. \end{aligned} \quad (\text{B.4})$$

The heaviside unit step function $\mathcal{H}(x - \xi)$ is defined to be zero if $x < \xi$, but defined to be one otherwise. The first term in the equation (B.4) applies if $L^* < L_c$, the remaining two terms if $L_c \leq L^*$. The second term represents the probability mass concentrated along the L^* axis due to the first segment of the reaction function (see the figure 1). Recalling that the Dirac delta function is the derivative of the heaviside unit step function, differentiation of the equation (B.4) with respect to both random variables yields the equation (2.2c) in the text.

Appendix C: The Optimality Conditions

In this appendix, we wish to demonstrate that the first-order optimality conditions are *not* tractable analytically even for the simplest case of a risk-neutral owner. Using the original probability density function, the objective function (2.7) can be transformed into the equation (C.1).

$$\begin{aligned}
 \mathcal{E}A &= \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=L_c/k(L^*)}^{\hat{D}=\infty} L^* f(\hat{D}, L^*) d\hat{D} dL^* & \text{①} \\
 &- L_c \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=L_c/k(L^*)}^{\hat{D}=\infty} f(\hat{D}, L^*) d\hat{D} dL^* & \text{②} \\
 &+ \alpha \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=0}^{\hat{D}=L_c/k(L^*)} L^* f(\hat{D}, L^*) d\hat{D} dL^* & \text{③} \\
 &- \alpha [L_c + S] \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=0}^{\hat{D}=L_c/k(L^*)} f(\hat{D}, L^*) d\hat{D} dL^* & \text{④} \\
 &+ [1 - \alpha] [r_c L - r_d D - \varphi(L, D, D+E-L)] & \text{⑤} \\
 &\cdot \int_{L^*=0}^{L^*=L_c} \int_{\hat{D}=L_c/k(L^*)}^{\hat{D}=\infty} f(\hat{D}, L^*) d\hat{D} dL^* \\
 &+ 0 \cdot \int_{L^*=0}^{L^*=L_c} \int_{\hat{D}=0}^{\hat{D}=L_c/k(L^*)} f(\hat{D}, L^*) d\hat{D} dL^*
 \end{aligned} \tag{C.1}$$

Here, we extend Leibniz' rule to a two-dimensional integral to be applied to the objective function above. Consider a two-dimensional integral with limits of integration which are dependent on a parameter ξ as follows:

$$\psi \equiv \int_{y=c(\xi)}^{y=d(\xi)} \int_{x=a(y;\xi)}^{x=b(y;\xi)} f(x, y; \xi) dx dy . \tag{C.2}$$

Observe that the region of integration is also bended in the y direction. The first derivative of this integral with respect to the parameter ξ is given by the extended Leibniz rule as follows

$$\begin{aligned}
 \frac{\partial \psi}{\partial \xi} &= \int_{y=c(\xi)}^{y=d(\xi)} \int_{x=a(y;\xi)}^{x=b(y;\xi)} \frac{\partial f(x, y; \xi)}{\partial \xi} dx dy \\
 &+ \int_{y=c(\xi)}^{y=d(\xi)} \frac{\partial b(y; \xi)}{\partial \xi} f(b(y; \xi), y; \xi) dy - \int_{y=c(\xi)}^{y=d(\xi)} \frac{\partial a(y; \xi)}{\partial \xi} f(a(y; \xi), y; \xi) dy \\
 &+ \frac{\partial d(\xi)}{\partial \xi} \int_{x=a(y;\xi)}^{x=b(y;\xi)} f(x, d(\xi); \xi) dx - \frac{\partial c(\xi)}{\partial \xi} \int_{x=a(y;\xi)}^{x=b(y;\xi)} f(x, c(\xi); \xi) dx .
 \end{aligned} \tag{C.3}$$

The objective function is broken down into the five terms indicated by the numbers ① – ⑤ in the equation (C.1) above due to its length. The first-order conditions for a maximum of

the expected final pay-offs are shown in the following equations for the dummy variable $\xi \in \{L, D, E\}$ given the fact that cash balances have been substituted from the balance sheet constraint at the beginning of the model period. The partial derivative of the first term is given by:

$$\begin{aligned} \frac{\partial \textcircled{1}}{\partial \xi} &= \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=L_c/\kappa(L^*)}^{\hat{D}=\infty} L^* \frac{\partial}{\partial \xi} f(\hat{D}, L^*) d\hat{D} dL^* - \frac{\partial L_c}{\partial \xi} \int_{\hat{D}=L_c/\kappa(L_c)}^{\hat{D}=\infty} L_c f(\hat{D}, L_c) d\hat{D} \\ &\quad - \int_{L^*=L_c}^{L^*=\infty} \left(\frac{\partial}{\partial \xi} \frac{L_c}{\kappa(L^*)} \right) L^* f\left(\frac{L_c}{\kappa(L^*)}, L^*\right) dL^* \end{aligned} \quad (\text{C.4})$$

The partial derivative of the second term is given by:

$$\begin{aligned} \frac{\partial \textcircled{2}}{\partial \xi} &= - \frac{\partial L_c}{\partial \xi} \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=L_c/\kappa(L^*)}^{\hat{D}=\infty} f(\hat{D}, L^*) d\hat{D} dL^* \\ &\quad - L_c \left\{ \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=L_c/\kappa(L^*)}^{\hat{D}=\infty} \frac{\partial}{\partial \xi} f(\hat{D}, L^*) d\hat{D} dL^* \right. \\ &\quad \left. - \int_{L^*=L_c}^{L^*=\infty} \left(\frac{\partial}{\partial \xi} \frac{L_c}{\kappa(L^*)} \right) f\left(\frac{L_c}{\kappa(L^*)}, L^*\right) dL^* - \frac{\partial L_c}{\partial \xi} \int_{\hat{D}=L_c/\kappa(L_c)}^{\hat{D}=\infty} f(\hat{D}, L_c) d\hat{D} \right\} \end{aligned} \quad (\text{C.5})$$

The partial derivative of the third term is given by:

$$\begin{aligned} \frac{\partial \textcircled{3}}{\partial \xi} &= \alpha \left\{ \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=0}^{\hat{D}=L_c/\kappa(L^*)} L^* \frac{\partial}{\partial \xi} f(\hat{D}, L^*) d\hat{D} dL^* \right. \\ &\quad \left. + \int_{L^*=L_c}^{L^*=\infty} \left(\frac{\partial}{\partial \xi} \frac{L_c}{\kappa(L^*)} \right) L^* f\left(\frac{L_c}{\kappa(L^*)}, L^*\right) dL^* - \frac{\partial L_c}{\partial \xi} \int_{\hat{D}=0}^{\hat{D}=L_c/\kappa(L_c)} L_c f(\hat{D}, L_c) d\hat{D} \right\} \end{aligned} \quad (\text{C.6})$$

The partial derivative of the fourth term is given by:

$$\begin{aligned} \frac{\partial \textcircled{4}}{\partial \xi} &= - \alpha \frac{\partial L_c}{\partial \xi} \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=0}^{\hat{D}=L_c/\kappa(L^*)} f(\hat{D}, L^*) d\hat{D} dL^* \\ &\quad - \alpha [L_c + S] \left\{ \int_{L^*=L_c}^{L^*=\infty} \int_{\hat{D}=0}^{\hat{D}=L_c/\kappa(L^*)} \frac{\partial}{\partial \xi} f(\hat{D}, L^*) d\hat{D} dL^* \right. \\ &\quad \left. + \int_{L^*=L_c}^{L^*=\infty} \left(\frac{\partial}{\partial \xi} \frac{L_c}{\kappa(L^*)} \right) f\left(\frac{L_c}{\kappa(L^*)}, L^*\right) dL^* - \frac{\partial L_c}{\partial \xi} \int_{\hat{D}=0}^{\hat{D}=L_c/\kappa(L_c)} f(\hat{D}, L_c) d\hat{D} \right\} \end{aligned} \quad (\text{C.7})$$

The partial derivative of the fifth term is given by:

$$\begin{aligned}
 \frac{\partial \textcircled{5}}{\partial \xi} = & [1 - \alpha] \frac{\partial B}{\partial \xi} \int_{L^*=0}^{L^*=L_c} \int_{\widehat{D}=L_c/\kappa(L^*)}^{\widehat{D}=\infty} f(\widehat{D}, L^*) d\widehat{D} dL^* \\
 & + [1 - \alpha] B(\cdot) \left\{ \int_{L^*=0}^{L^*=L_c} \int_{\widehat{D}=L_c/\kappa(L^*)}^{\widehat{D}=\infty} \frac{\partial}{\partial \xi} f(\widehat{D}, L^*) d\widehat{D} dL^* \right. \\
 & \left. - \int_{L^*=0}^{L^*=L_c} \left(\frac{\partial}{\partial \xi} \frac{L_c}{\kappa(L^*)} \right) f\left(\frac{L_c}{\kappa(L^*)}, L^* \right) dL^* + \frac{\partial L_c}{\partial \xi} \int_{\widehat{D}=L_c/\kappa(L_c)}^{\widehat{D}=\infty} f(\widehat{D}, L_c) d\widehat{D} \right\} \quad (\text{C.8}) \\
 B = & r_e L - r_d D - \varphi(L, D, D+E-L)
 \end{aligned}$$

The first-order optimality conditions (C.4) – (C.8) form a non-linear equation system involving two-dimensional integrals for loans, deposits and equity. However, if we restrict the total balances of our bank to be one, either deposits or equity can be eliminated further. The remaining equation system is not tractable analytically. Although our model is simple, only a numerical sensitivity analysis can unveil the optimal equity-to-debt ratio. Moreover, any comparative static analysis seems to be impossible.

Appendix D: Numerical Quadrature in Two Dimensions

In this appendix, we describe the key operation of the numerical sensitivity analysis, namely, how to integrate the objective function (2.7). We suggest to apply the trapezoidal rule as described by Waldvogel [1988]. In the case of an analytic integrand, this operation is computationally much more efficient than the ROMBERG integration algorithm which is highly recommended by Press et al. [1989, section 4.3]. The use of the trapezoidal rule for analytic integrands has been suggested earlier by several authors, see e. g. Schwartz [1969] or Iri et al. [1970], but the method entered the textbooks only recently, see e. g. Schwarz [1986]. For instance, Waldvogel's algorithm achieves machine precision for the probability integrals over the whole range of definition for the univariate normal density function and the univariate non-central chi-square density function, whereas the ROMBERG algorithm fails.

The numerical quadrature algorithm considers three standard types of ranges of integration, namely $\mathcal{F}_1 \equiv (-\infty, +\infty) \times (-\infty, +\infty)$, $\mathcal{F}_2 \equiv (0, +\infty) \times (0, +\infty)$ and $\mathcal{F}_3 \equiv (-1, +1) \times (-1, +1)$. Of course, any integral with a rectangular range of integration can be either transformed into an integral with one of these three ranges by means of a linear transformation or represented as a combination of the three standard types. Suppose that the integral $\mathfrak{S} := \int \int_{t \in \mathcal{F}} f(t) dt$ exists and that $f(t) \rightarrow 0$ as $t \rightarrow \pm \infty$ if the limits of integration are infinite. Integrable singularities can be treated at the limits of integration in the case of the single-infinite range $\mathcal{F} = \mathcal{F}_2$ or the finite range $\mathcal{F} = \mathcal{F}_3$. Each of the three types of integrals is transformed into an improper integral with both limits unbounded, that is

$$\mathfrak{S} := \iint_{x_{n+2}, y_{n+2} \in \mathcal{F}} f(x_{n+2}, y_{n+2}) dy_{n+2} dx_{n+2} = \int_{x_0=-\infty}^{x_0=+\infty} dx_0 \int_{y_0=-\infty}^{y_0=+\infty} f(x_0, y_0) p_x p_y dy_0 \quad (\text{D.1})$$

This is exactly the reverse strategy of what is usually done, see e. g. Press et al. [1989]. The transformation of variables in the above equation consists of n optional transformations with the sinh-function, one rotation of the axes, and one compulsory transformation with a particular function which is suited for each of the three ranges of integration:

$$\left\{ \begin{array}{l} x_j = \sinh(x_{j-1}) \\ y_j = \sinh(y_{j-1}) \\ (j = 1, \dots, n) \end{array} \right\}, \quad \left\{ \begin{array}{l} x_{n+1} = \cos(\varphi) x_n - \sin(\varphi) y_n \\ y_{n+1} = \sin(\varphi) x_n + \cos(\varphi) y_n \end{array} \right\}, \quad \left\{ \begin{array}{l} x_{n+2} \\ y_{n+2} \end{array} \right\} = \left\{ \begin{array}{ll} \left\{ \begin{array}{l} x_{n+1} \\ y_{n+1} \end{array} \right\} & \text{if } \mathcal{F} = \mathcal{F}_1, \\ \left\{ \begin{array}{l} \exp(x_{n+1}) \\ \exp(y_{n+1}) \end{array} \right\} & \text{if } \mathcal{F} = \mathcal{F}_2, \\ \left\{ \begin{array}{l} \tanh(x_{n+1}) \\ \tanh(y_{n+1}) \end{array} \right\} & \text{if } \mathcal{F} = \mathcal{F}_3. \end{array} \right. \quad (\text{D.2})$$

In the case of the first type of integral with range $\mathcal{F} = \mathcal{F}_1$, the compulsory transformation is, of course, just the identity transformation because the original integral is already improper. Since the bivariate lognormal probability density function stretches along the 45° line for high coeffi-

icients of correlation in absolute value, the angle of rotation, φ , may be chosen to be $\pm 45^\circ$ for $|\rho| \geq 0.5$, say. The absolute value of the Wronsky determinant is equal to $p_x p_y$. The term p_x may be written as

$$p_x = \begin{cases} \prod_{j=0}^{n-1} \cosh(x_j) & \text{if } \mathcal{F} = \mathcal{F}_1, \\ e^{x_{n+1}} \prod_{j=0}^{n-1} \cosh(x_j) & \text{if } \mathcal{F} = \mathcal{F}_2, \\ \frac{1}{\cosh^2(x_{n+1})} \prod_{j=0}^{n-1} \cosh(x_j) & \text{if } \mathcal{F} = \mathcal{F}_3. \end{cases} \quad (\text{D.3})$$

Replace x with y to get p_y . The integrals of the objective function (2.7) are now evaluated by means of the simplest quadrature scheme, that is, the trapezoidal rule. Given a choice of initial step sizes $h = (h_x, h_y)$, the function values, which are evaluated at equally spaced abscissas, are summed in both the positive and negative directions of integration. Given a center of integration $c = (c_x, c_y)$ as the starting point, the summation is stopped as soon as the partial sum does not change any more:

$$T(h) = h_x h_y S(h), \quad S(h) \equiv \sum_{x_i = c_x \pm i h_x}^{|g(\cdot)| < \varepsilon} \sum_{y_j = c_y \pm j h_y}^{|g(\cdot)| < \varepsilon} g(x_i, y_j), \quad g(\cdot) \equiv f(\cdot) p_x(\cdot) p_y(\cdot), \quad (\text{D.4})$$

where $\varepsilon = 1.0 \cdot 10^{-19}$ is the machine tolerance (of the Apple Macintosh personal computer). Next, the step sizes are cut in half and the new center of integration is equal to the old center shifted by the new step sizes. Then the summation is repeated in all directions as in the equation (D.4). Next, the step sizes are cut in half again, the summation is continued, and so forth. The repeated reduction of the step sizes is stopped when $|T(h) - T(h/2)| < \sqrt{\varepsilon}$. Then, $T(h/2)$ has accuracy ε . The convergence of the trapezoidal value to the integral is given by $T(h) - \mathfrak{I} = \mathcal{O}(e^{-\gamma/h})$ with γ a positive constant, given an analytic function $g(\cdot)$.

For our numerical sensitivity analysis, we used one sinh-transformation. Numerical experiments show that the trapezoidal rule is generally quite efficient if the decay of the integrand $g(x_0, y_0)$ as $(x_0, y_0) \rightarrow \pm \infty$ is doubly exponential (see Takahashi and Mori [1974]).

References

- Arrow, Kenneth J. [1971]: *Essays in the Theory of Risk-Bearing*, Chicago: Markham Publishing Company.
- Baltensperger, Ernst [1980]: “Alternative Approaches to the Theory of the Banking Firm”, *The Journal of Monetary Economics*, 6: 1 – 37.
- Baltensperger, Ernst and Hellmuth Milde [1987]: *Theorie des Bankverhaltens*, Springer-Verlag: Heidelberg.
- Büttler, Hans-Jürg and Kurt Schiltknecht [1983]: “Transitory Changes in Monetary Policy and Their Implications in Money Stock Control”, in K. Brunner and A. Meltzer (eds.): *Carnegie-Rochester Conference Series on Public Policy*, 19: 171 – 210, Amsterdam: North-Holland Publishing Company.
- Büttler, Hans-Jürg and Jörg Waldvogel [1996]: “Pricing Callable Bonds by Means of Green’s Function”, *Mathematical Finance*, 6: 53 – 88.
- Calomiris, Charles W. and Charles M. Kahn [1991]: “The Role of Demandable Debt in Structuring Optimal Banking Arrangements”, *The American Economic Review*, 81: 497 – 513.
- Char, V. V. and Ravi Jagannathan [1988]: “Banking Panics, Information, and Rational Expectations Equilibrium”, *The Journal of Finance*, 63: 749 – 763.
- Diamond, Douglas W. and Philip H. Dybvig [1983]: “Bank Runs, Deposit Insurance, and Liquidity”, *The Journal of Political Economy*, 91: 401 – 419.
- Fama, Eugene F. [1980]: “Banking in the Theory of Finance”, *The Journal of Monetary Economics*, 6: 39 – 57.
- Fama, Eugene F. [1985]: “What’s Different About Banks?”, *The Journal of Monetary Economics*, 15: 29 – 39.
- Gorton, Gary [1985]: “Bank Suspension of Convertibility”, *The Journal of Monetary Economics*, 15: 177 – 193.
- Haugen, Robert A. and Lemma W. Senbet [1978]: “The Insignificance of Bankruptcy Costs to the Theory of Optimal Capital Structure”, *The Journal of Finance*, 33: 383 – 393.
- Iri, M., S. Moriguti, and Y. Takasawa [1970]: “On a certain Quadrature Formula”, *RIMS Kokyuroku Kyoto University*, 91: 82 – 118.
- Jacklin, Charles J. and Sudipto Bhattacharya [1988]: “Distinguishing Panics and Information-based Bank Runs: Welfare and Policy Implications”, *The Journal of Political Economy*, 96: 568 – 592.
- James, Christopher [1991]: “The Losses Realized in Bank Failures”, *The Journal of Finance*, 46: 1223 – 1242.
- Lacker, Jeffrey M. [1991]: “Why Is There Debt?” *Economic Review*, Federal Reserve Bank of Richmond, July/August, 3 – 19.
- Merton, Robert C. [1971]: “Optimum Consumption and Portfolio Rules in a Continuous-Time Model”, *The Journal of Economic Theory*, 3: 373 – 413.
- Miller, Merton H. [1977]: “Debt and Taxes”, *The Journal of Finance*, 261 – 75.
- Miller, Merton H. [1988]: “The Modigliani-Miller Propositions After Thirty Years”, *The Journal of Economic Perspectives*, 2 (4) : 99 – 120.
- McCulloch, Huston and Min-Teh Yu [1993]: “Government Deposit Insurance and the Diamond-Dybvig Model”, working paper, Department of Economics, Ohio State University.
- Modigliani, Franco and Merton H. Miller [1958]: “The Cost of Capital, Corporation Finance and the Theory of Investment”, *The American Economic Review*, June, 48: 261 – 97.
- Modigliani, Franco [1988]: “MM – Past, Present, and Future”, *The Journal of Economic Perspective*, 2 : 149 – 158.
- Orgler, Yair E. and Robert A. Taggart [1983]: “Implications of Corporate Capital Structure Theory for Banking Institutions”, *The Journal of Money, Credit, and Banking*, 15: 212 – 21.
- Press, William H., Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling [1989]: *Numerical Recipes in Pascal: The Art of Scientific Computing*, Cambridge (Mass.): Cambridge University Press.
- Pringle, John J. [1974]: “The Capital Decision in Commercial Banks”, *The Journal of Finance*, 58: 779 – 95.
- Russell, Steven [1993]: “Comments on ‘Government Deposit Insurance in the Diamond-Dybvig Model’ “ by J. Huston McCulloch and Min-Teh Yu, working paper, Department of Economics, Ohio State University.
- Santomero, Anthony M. [1984]: “Modelling the Banking Firm”, *The Journal of Money, Credit, and Banking*, 16: 576 – 616.
- Schwartz, C. [1969]: “Numerical Integration of Analytic Functions”, *The Journal of Computational Physics*, 4: 19 – 29.
- Schwarz, Hans Rudolf [1986]: *Numerische Mathematik*, Stuttgart: B. G. Teubner, 2nd edition.
- Takahashi, H. and M. Mori [1974]: “Double Exponential Formulas for Numerical Integration”, *Publications RIMS Kokyuroku Kyoto University*, 9: 721 – 741.
- Talmor, Eli [1980]: “A Normative Approach to Bank Capital Adequacy”, *The Journal of Financial and Quantitative Analysis*, November, 14: 785 – 811
- Waldvogel, Jörg [1988]: “Numerical Quadrature in Several Dimensions”, *International Series of Numerical Mathematics*, 85: 295 – 309, Basel: Birkhäuser Verlag.
- Warner, J. [1977]: “Bankruptcy Costs: Some Evidence”, *The Journal of Finance*, May, 56: 337 – 47.