Banking Scope and Financial Innovation

Arnoud W. A. Boot
University of Amsterdam

Anjan V. Thakor
University of Michigan

We explore the implications of financial system design for financial innovation. We begin with assumptions about the investment opportunities of firms, their observable attributes, and the roles of commercial banks, investment banks, and the financial market. We examine the borrower's choice between bank and financial market funding, the commercial bank's choice of monitoring capacity, and the investment bank's choice of whether to invest in financial innovation. Our main result is that financial innovation in a universal banking system is stochastically lower than innovation in a financial system in which commercial and investment banks are functionally separated.

Perhaps it is this specter that most haunts the working men and women: the planned obsolescence of people that is of a piece with the planned obsolescence of the things they make. — Studs Terkel

---

The authors thank Anjolein Schmeits (in particular) and Kathleen Petrie for research assistance. For their many helpful comments, we gratefully acknowledge Franklin Allen (the editor), an anonymous referee, Andy Winton (discussant at the WFA-RFS Symposium), and participants at the NYU/Salomon Brothers Conference on Universal Banking (1996) and the Western Finance Association Meeting in Sunriver (1996), as well as participants at finance workshops at the University of Miami, University of Utah, University of Iowa, Indiana University, London School of Economics, University of Lausanne, Universitat Pompeu Fabra (Barcelona), Univeristy of Amsterdam, and Tilburg University for their helpful comments. We alone are responsible for the contents of this article. Address correspondence to Anjan V. Thakor, Edward J. Frey Professor of Banking and Finance, University of Michigan Business School, 701 Tappan Street, Ann Arbor, MI 48109.

The Review of Financial Studies Winter 1997 Vol. 10, No. 4, pp. 1099–1131

We study the implications of alternative designs of the financial system with a view to improving our understanding of the pros and cons of functionally separated banking (the U.S. system, for example) vis-à-vis universal banking (the German system, for example). There has been a great deal of practical interest in this subject as exemplified by the following quote from *The Economist* (1994):

> What do the Porsche 911 and Deutsche Bank have in common? The answer is that both these German creations are widely considered to be perfect models — and nowhere more so than in Central Europe. While car lovers around the world admire the Porsche's sleek lines, bankers and policy makers in Warsaw, Prague and Budapest are impressed by lines of another kind: those on Deutsche's balance sheet.... This model of "universal" banking has sometimes been seen as a cornerstone of Germany's post-war economic success. Unsurprisingly, neighboring countries that are rebuilding their financial systems from the rubble of communism are tempted to copy it. That would be a mistake.

Academic research has kept abreast of the practical interest in this topic. There are three strands of the literature that are relevant. First is the research on financial innovation and security design [see, e.g., Allen and Gale (1988, 1991, 1994), Bhattacharya and Nanda (1996), Boot and Thakor (1993), Duffie and Rahi (1995), Nanda and Yun (1994), and Riddiough (1997)]. This literature explains what motivates financial innovation and how securities are designed, priced, and marketed. A second literature — that has grown somewhat independently — is concerned with the policy question of banking scope, that is, whether the banking system should contain functionally separated commercial and investment banks or universal banks [see, e.g., Berlin, John, and Saunders (1994), Kanatas and Qi (1994), Puri (1994), Kroszner and Rajan (1994a, 1994b), and Rajan (1993)]. The focus here has largely been on potential conflicts of interest associated with universal banking. Somewhat more recently, attention has focused on the broader issue of *financial system design* [see, e.g., Allen (1992), Allen and Gale (1995), Boot and Thakor (1997), Neave and Johnson (1993), Sabani (1992), and Titman and Subrahmanyam (1996)]. This literature has addressed a comprehensive set of questions concerned with how financial system design affects individual risk-sharing opportunities, the allocation and cost of capital for corporations, corporate governance, and the restructuring of firms in financial distress. Since the design of contracts, institutions, and markets, as well the determination of banking scope, are all part of the details of how a financial system should be configured, the emerging literature on financial system design promises to provide valuable unifying insights.

This article focuses on the effect of financial system design on financial innovation. In particular, we examine the impact of banking scope — the choice between universal and functionally separated banking — on the endogenously determined incentives of institutions to engage in financial innovation, and thus on each borrower's choice of financing source and its cost of capital. In addition to explaining how financial innovation is influenced by banking scope, the analysis speaks to a host of related system design issues, such as the implications of banking industry fragmentation or consolidation, the potential path dependence in the evolution of the financial system, and the desirable starting point of a new financial system. Thus our research touches all three strands of the literature mentioned earlier — financial innovation and security design, the implications of banking scope, and overall financial system design and evolution.

The model is characterized by four key players: commercial banks, investment banks, borrowing firms, and the financial market. The actions of each are endogenously determined according to an optimization program. Commercial banks specialize in postlending monitoring to deter asset-substitution moral hazard. Investment banks assist borrowers in raising funds in the capital market and design securities (through financial innovation) to lower their borrowers' cost of capital. Borrowers optimize through their choice of financing source, which is predicated on an observable attribute that varies cross-sectionally. The financial market consists of informed and other traders. How many traders become informed (and hence trading volume) depends on the design of securities and the attributes of firms that access the capital market. Thus the actions of investment banks and borrowers impact the "price efficiency" of the capital market. The advantage of capital market financing for the borrower is that informed traders possess payoff-relevant information that the borrower does not have and this information is noisily transmitted to the borrower through the market price of its debt security, thereby leading to improved real decisions and an enhanced payoff.

In this setting, the borrower trades off the advantage of bank financing (which lies in the bank's ability to attenuate asset-substitution moral hazard) against the advantage of capital-market financing (which stems from the feedback role of capital market prices).[1] We assume that the severity of the borrower's moral hazard is captured by a publicly observable quality attribute, with lower values of this attribute

[1] Thus our analysis sidesteps the issue of the relationship between the borrower's choice of financing source and the extent to which there is leakage of proprietary information (that the borrower does not wish to disclose) to the borrower's competitors due to the process of raising financing. These issues are examined in Bhattacharya and Chiesa (1995) and Yosha (1995).

representing more severe moral hazard. It can then be shown that there is a "quality cutoff" in the borrower's choice of financing source. Borrowers below this quality cutoff approach banks because the moral hazard problem is the most severe for them, whereas borrowers above this cutoff access the capital market. Since this cutoff is endogenously determined by the tension faced by the borrower between the value of moral hazard amelioration and the value of the information conveyed by the capital market price, financial innovation affects this cutoff as well. If an investment bank can design a *new* security that results in the equilibrium security price reflecting more of the information possessed by the informed agents, then this innovation will cause the quality cutoff to decline as more borrowers gravitate to the capital market.

If the financial system has functional separation between commercial and investment banks, then each investment bank will choose its investment in innovation based on the cost of the innovation relative to the expected increase in its fee revenue that comes from sharing in the borrower's elevated payoff due to the innovation. But the decision rule is different if we have universal banking. Now the investment banking arm of the universal bank internalizes the potentially pernicious effect of financial innovation on the customer base of the commercial banking arm, that is, the commercial bank's borrowers may defect to the efficiency-enhanced financial market. The equilibrium level of financial innovation is lowered as a consequence. This provides one perspective on the higher rate of financial innovation in the U.S. relative to Europe.

The structure of the banking industry, manifested in its fragmentation/competitiveness, affects *interbank* competition and hence the price at which commercial bank credit is available. This leads to a link between banking industry structure, either with functional separation or with universal banking, and the quality cutoff that delineates bank borrowers from capital market borrowers (even ignoring financial innovation incentives). Moreover, the sophistication of the financial market is an important determinant of the impact of a financial innovation. For example, the introduction of an exotic new option is likely to be less successful in an underdeveloped financial market than in a more developed, sophisticated financial market. But the success or failure of the financial innovation in turn affects the *future* evolution of the financial market. Hence the evolution of the financial market is likely to be path-dependent [see also Dinc (1994)].

Our analysis points, therefore, to the many important effects that financial system design is likely to have on credit allocation and economic development. The ramifications of this for the structuring of financial systems in ex-communist economies are transparent and

echoed in the following quote from *The Economist* (1994):

> Yet the German model may not be suitable for economies that are making the painful transition from central planning to capitalism. One priority should be to create a stable banking system that wins depositors' trust while allocating credit on the basis of market forces. A second should be to encourage a rapid restructuring of the hugely inefficient industries that central planning has created. And a third should be to promote the development of efficient and competitive capital markets. An unthinking dash for a universal-banking system could make it harder to meet any of these priorities.

We have focused in our analysis on the impact of two key aspects of financial system design on financial innovation: the degree to which the banking system is functionally separated (or universal) and the degree of fragmentation in the banking system. Both aspects are important in driving our results. In particular, the deleterious effect of universal banking on financial innovation predicted by our analysis presupposes a high degree of consolidation with universal banks. Without such consolidation, a universal bank would not discern a dampening of the demand for its loans due to its own financial innovation. Thus, in a very fragmented universal banking system — perhaps like that which existed in the United States prior to the Glass Steagall Act — financial innovation would not be significantly discouraged by the universal nature of banking.

The rest is organized as follows. Section 1 presents the model. Section 2 presents an analysis of the borrower's choice of financing source. Section 3 contains an analysis of the decisions of commercial and investment banks for a financial system with functionally separated commercial and investment banking as well as for a financial system with universal banking. Section 4 discusses key implications. Section 5 concludes. All proofs are in the Appendix at the end of the article.

1. The Model

1.1 Investment choices of firms

There is universal risk neutrality, and the riskless rate is zero. Each firm in the economy has the potential to invest in a single-period project that needs a \$1 investment. Whether the project will actually become available to a borrower one period hence is uncertain at the outset; this uncertainty will be resolved at $t = 1$. Moreover, conditional on a project being available, the quality of the project is random. Conditional on an investment opportunity being available, the probability is $\theta \in (0, 1)$ that the firm has only a good project available. This

project yields a terminal payoff of $\$Y > 0$ with probability (w.p.) $\eta \in (0, 1)$ and 0 w.p. $1-\eta$. With probability $1-\theta$, the firm will have a choice between this good project and a bad project. The latter yields a contractible payoff of 0, but generates a noncontractible private rent $R > 0$ for the firm's manager; this could be viewed as a control rent as in O'Hara (1993). We will later impose parametric restrictions that ensure that the manager will always prefer the bad project with external financing even though he would prefer the good project if he could self-finance (the firm's cash constraints preclude self-financing).

Each potential borrower is characterized by an observable parameter $\theta \in (0, 1)$. Each borrower knows its own θ at the outset, but others observe it only a period later. Let G be the cumulative distribution over the cross-section of θ's, and $g(\theta)$ the associated density function that outsiders associate with θ. This parameter θ is the commonly known prior probability assigned by the market to the event that a borrower with that θ will have access only to the good project, and therefore pose no asset-substitution moral hazard problem.

1.2 Role of commercial banks

Commercial banks (CBs) specialize in postlending monitoring that resolves asset-substitution moral hazard. Thus, if a firm borrows from a bank, the choice of the good project can be ensured w.p. 1. The bank incurs a cost $C > 0$ to monitor each borrower, and it must decide at the outset how much monitoring capacity to acquire for the period. Let N_0 denote the monitoring capacity the bank acquires at the beginning of the period, at a total cost of CN_0. With this capacity, the bank can monitor at the most N_0 borrowers. The loan demand the bank faces is random, however. The realized loan demand for a bank depends on numerous factors, including the number of borrowers who will need funds, the realizations of θ, and the decisions of borrowers about whether to go to banks or the capital market. If loan demand exceeds N_0, then the demand in excess of N_0 must either be rationed or extended loans without postlending monitoring. If loan demand falls short of N_0, then the excess of N_0 over the realized loan demand remains unutilized.

We visualize an imperfectly competitive banking industry. As in Besanko and Thakor (1992), we can imagine banks lying along the circumference of a circle, engaging in competition constrained by spatial considerations on the part of borrowers.[2] In particular, we view the lack of perfect competition — and any rents arising therefrom — as

[2] The spatial representation is best viewed as an allegory for more general product-differentiation-based imperfections in competition.

related to the bank's monitoring ability. Each bank realizes a particular (random) loan demand in its "area" and "transportation costs," for borrowers permit the bank to earn a rent of $\tau > C$ on each borrower it monitors, that is, the bank earns a rent that is such that the borrower is indifferent between paying that rent and incurring the cost to go to the next most convenient bank.

We can view τ as a rent that compensates the bank for some previously incurred fixed cost of entry into the banking industry and having acquired the expertise to monitor borrowers at a cost. Another interpretation is that these rents arise from ex post informational monopolies stemming from information flows in durable bank-borrower relationships [e.g., Rajan (1992)]. Clearly the assumption that this rent is fixed per borrower is strong. In general, we would expect the rent to arise from an explicit consideration of the borrower's various financing alternatives, in which case it is likely that τ would be decreasing in θ since the bank would have to settle for a lower rent from higher quality borrowers who have a lower "need" for bank monitoring. This will not affect the qualitative nature of our results. In any case, our goal is not to explain the existence of banking rents, but rather the implications of these rents for lending policy and innovation incentives. Nonetheless, a more fully developed model in which these rents arise endogenously would be interesting.

One approach to building such a model is suggested by the work of Kreps and Scheinkman (1983) on the relationship between Bertrand and Cournout competition. They show that to reach the perfectly competitive outcome that is usually associated with Bertrand competition requires *both* the assumption that firms compete on prices and the assumption that production occurs after demand is determined in response to the announced prices. If one assumes a two-stage game in which in the first stage firms determine their production capacities independently and simultaneously, then produce and bring these quantities to the market, and in the second stage engage in Bertrand-like price competition, then it is possible for the Cournot outcome to be the unique equilibrium. This kind of two-stage game fits our commercial banking system nicely since banks are first building up monitoring capacities and then competing (imperfectly) for borrowers. The Kreps and Scheinkman (1983) results indicate that banks can earn rents (like Cournot oligopolists) even when they are engaged in Bertrand competition of this sort.[3]

[3] Proceeding formally along the lines of Kreps and Scheinkman (1983) in our context is complicated, however, because loan demand is random and exogenously specified for each θ in our model for the set of θ's that approach banks, whereas they consider a deterministic, downward-sloping, price-dependent demand schedule. Endogenizing loan demand would significantly complicate

1.3 The capital market

The basic idea we want to model is that the capital market includes traders who acquire costly information relevant to the real decisions of firms that even the managers of these firms may not possess. This perspective differs from traditional signaling models in finance where the firm's manager is the one endowed with propreitary information. While we do not dispute the assumption that managers often know more about their firms than anyone else, we also believe that there are situations in which managers could learn something of value from the market. For instance, we could envision traders/analysts who are industry specialists who develop special skills in assessing shifts in customer preferences, changes in the competitive structure of the industry, and so on.

These security analysts may acquire privileged information randomly. For example, a security analyst who specializes in the pharmaceutical industry may learn something that may be of value to Eli Lilly's managers in a particular period. The analyst may or may not be able to credibly communicate to Lilly management that he has proprietary information that they should pay for; Fishman and Hagerty (1995) explain why informed traders may wish to sell their information, and Allen (1990) analyzes the credibility problem in direct information sales. In the next period, there may be a different analyst who acquires proprietary information. In a market with many analysts following the industry, it would be difficult for Lilly to ascertain who knows what and when. This would preclude Lilly from going out and hiring these traders to acquire their information directly.[4] Alternatively, even if no individual analyst/investor is better informed than the managers, it is nevertheless possible for the capital market in the aggregate to be better informed. For example, if individual traders who invest in information were to receive identical and independently distributed (noisy) signals, the market price could aggregate their information and reveal something that the firm's management was unaware of.[5]

---

our analysis since then τ would depend on the *number* of competing banks which would then also affect the quality (θ) cutoff determining which borrowers go to banks and which go to the capital market. We do not believe this will change our qualitative results, however.

[4] An exception to this would be when a significant portion of the firm's stock is owned by institutional investors who are informed. In this case, management could directly ask these investors for input. Pound (1997) describes a meeting convened by NewTech Corporation to which the company's largest institutional investors were invited to provide input that would help the company formulate strategies.

[5] The notion that there may be those in the market who know something of decision relevance that the firm's managers don't is evidenced by the NewTech Corporation case described by Pound (1997). We quote from Pound:

> In July of 1995, an unusual meeting took place at the Intercontinental Hotel in New York. It was convened by NewTech Corporation — a growing, successful public company, operating in two broad areas of applied information technology.... The subject of the meeting was of

The key to the information-feedback role of prices is that the informed traders will attempt to profit from their information by taking positions in the securities issued by the firms about which they have superior information. Although the presence of liquidity-motivated trades will mask the trades of the informed traders, the total order flow will at least noisily reveal informed trading. Based on this, the firm may be able to infer some of the information possessed by the informed traders and this may lead it to make a value-enhancing real decision.[6] This is one way to visualize the information aggregation role of the capital market and the feedback role of prices. The informed traders observe a market opportunity that they conjecture the firm will exploit and thus take a position in the firm's securities based on that conjecture, and the firm noisily infers the availability of this opportunity from the order flow for its securities and acts on it, thereby rationalizing the initial conjecture. We will now formalize this intuition about the interaction between the real decisions of firms and capital market price determination.

Suppose there are two types of investors/traders in the capital market: liquidity traders and discretionary agents. The aggregate demand of the liquidity traders for any asset is random and exogenously specified. A discretionary agent can become an "informed" agent at a private cost. This investment generates a privately observed signal, ϕ, that reveals payoff-relevant information about the firm's operating environment. Each informed agent receives exactly the same signal. This information can be "favorable" (f) or "unfavorable" (u). If $\phi = f$,

---

broad and fundamental importance to NewTech. The Company's officers had gathered for an all-day retreat to assess the Corporation's overall corporate structure, strategy, and prospects in the next five years.... The unusual aspect of the NewTech meeting was not the decision to hold such an offsite strategic review. Instead, the remarkable part of the NewTech meeting was the presence around the conference table of ten additional individuals. These individuals were not members of NewTech's management or its board; nor did they represent the usual coterie of paid corporate advisers. Instead, they represented ten of NewTech's largest and best-informed institutional shareholders.... Now, NewTech was faced with developing a defining strategy that would guide corporate policy going forward. This would require not execution, but creative insight — vision.... That, Strideman emphasized was the key reason that NewTech's top institutional investors had been invited to sit around the table.... Who better to judge the Company's initiatives in the context of what its competitors were doing — firms that the institutional investors in the room also followed and had detailed knowledge of? ... Strideman emphasized that the purpose of the day's proceedings was thus not to give shareholders any new quantitative information about NewTech. Instead, the purpose of the day was for NewTech to get information from *them*.

Clearly, direct communication like this is possible if *all* investors with valuable information can be identified ex ante. If this is not possible, the next best alternative may be to infer what they know from market prices.

[6] Consider the following example of a value-enhancing decision. Suppose we have one company producing VHS VCRs and another producing BETA VCRs, and each must decide how much to invest in expanding productive capacity. Whether this investment is good for the company's shareholders depends on which "standard" will ultimately prevail. Each company's stock price will reveal the aggregated information of the market about whether BETA or VHS will become the standard, and this could in turn guide each company's investment decision.

then the firm can make real investment decisions that can enhance its good project's payoff to $Y + \alpha$ w.p. η and α w.p. $1 - \eta$, where $\alpha \in (0, 1)$; the cost of this payoff-enhancing investment is $K > 0$. If $\phi = u$, then the payoff enhancement opportunity does not exist. This signal ϕ is unavailable to the firm's manager, but if the informed agents demand the security only when $\phi = f$, then the manager can infer valuable information from the aggregate demand for the security or its price. This inference will be noisy, however, because of liquidity trade randomness. For a similar approach to modeling the real impact of the capital market, see Allen (1992), Boot and Thakor (1996), and Holmstrom and Tirole (1993).[7]

The larger the fraction of the total trade volume that is potentially accounted for by informed traders, the more revealing is the order flow, and the smaller is the expected gain to each informed trader from his information. Thus the measure of informed traders, Ω, is endogenously determined through an equilibrium condition which states that the equilibrium value of Ω should be such that each discretionary agent is indifferent between becoming informed and staying uninformed, that is, the expected profit of each informed agent, net of the cost of becoming informed, should be zero.

The equilibrium price of the security is set to be equal to its expected value, with the expectation conditioned on the information contained in the aggregate demand, D, for the security; thus the discretionary uninformed traders earn zero expected profit on their trading. One can think of a competitive market maker setting the equilibrium price to clear the market, after observing D but being unable to distinguish the individual components of the demand attributable to the different types of traders. We also assume no short sales by agents other than the market maker and that the market maker absorbs any supply/demand imbalances.

The capital market has no monitoring capability. Thus, if the firm has a project available and further has a choice of project (conditional on a project being available, the probability of this is $1 - \theta$), it is anticipated that the bad project will be chosen by the manager. The market maker takes this into account in setting the security price. Moreover, she also accounts for the fact that there are some (sufficiently high) values of D such that project-payoff enhancement will occur and other (sufficiently low) values of D for which it will not. To

[7] We are assuming in our formal modeling that the prospects of firms seeking external financing are completely idiosyncratic. In reality, this will not always be true, so that even a firm that opts for a bank loan will be able to learn something from the prices of other firms' securities issued by investment banks. This will reduce the value of capital market financing relative to bank financing. We thank the referee for this observation.

ensure comparability with the bank financing case, we assume that capital market funds are raised through debt securities. It should be noted, however, that the type of security used for financing does not affect the analysis. In particular, the asset-substitution moral hazard problem here cannot be resolved by using equity instead of debt. The borrower always prefers the bad project with external financing, and only bank monitoring can ensure selection of the good project.

1.4 The role of the investment bank

The investment bank's (IB's) role is to underwrite the firm's debt offering in the capital market. Moreover, the IB can engage in security design innovation that improves the information sensitivity of the securities offered by the firm, as in the model developed by Boot and Thakor (1993). In the formal Boot–Thakor analysis, financial innovation takes the form of splitting a composite cash-flow security into debt and equity. This splitting creates one security (equity) that is more information-sensitive than the original composite security and one (debt) that is less information sensitive than the composite security. Traders are wealth-constrained and therefore have limited wealth to invest in securities. Financial innovation permits informed traders to devote their entire investment to the most information-sensitive security and thus increases their marginal return on investment in costly information. This induces more traders to become informed, leading to a higher endogenously determined measure of informed traders in equilibrium. As a consequence, order flow becomes more informative for the market maker and the good (undervalued) firm is able to realize a higher expected *total* equilibrium revenue for its debt and equity issues[8] than it could when it issued a single composite security.

In our analysis, we simplify by shying away from the details of financial innovation, viewing it instead as the creation of an unspecified feature in the design of the debt contract — for example, an option or a callability feature — that makes that security more information sensitive, and this induces more informed trading. This benefits the firm in two ways when it invests in the good project. First, it improves the information content of D, and thus leads to a higher probability of realizing the project payoff enhancement. Second, the higher probability of payoff enhancement reduces the expected cost of borrowing in the capital market.[9]

[8] The security splitting in Boot and Thakor (1993) creates a riskless debt security for which the firm receives a first-best price.

[9] When the informed bid for the security and D is high enough to convince the firm to take advantage of the opportunity to enhance the good project's payoff, the payoff to bondholders increases by α in the state in which it would be 0 without the enhancement initiative. This lowers the interest rate the firm must pay.

To come up with the financial innovation, the IB must understand the borrower's idiosyncratic project. This means that the IB must invest some resources to study the borrower *before* it knows whether it will get the borrower's business. Thus the manner in which IBs compete is as follows. At $t = 0$, each IB decides which borrower it will study. To ensure a competitive market, we assume that borrowers are scarce relative to IBs, so that each borrower may have multiple IBs studying it. After an IB knows how many other IBs it is competing with, it innovates with an endogenously determined innovation probability. If it innovates, it will make an investment of $\xi > 0$, which covers its cost of studying the firm and innovating. This investment results in a successful innovation with probability one. It does not matter to the analysis if we make the outcome of the innovation initiative random.

We assume that the IB captures the entire gain to the issuing firm from financial innovation through an increase in its fee, as long as it faces no competition from other IBs. This is possible only if the IB is the only institution that comes up with the innovation. If there is another IB that comes up with the same innovation, then none can profit from the innovation because they compete away their rents through a standard Bertrand undercutting argument. Moreover, none can recoup ξ.

1.5 Sequence of events

There are three dates: $t = 0, 1, 2$. At $t = 0$, each commercial bank chooses its monitoring capacity N_0, and each investment bank determines the probability with which it will invest in financial innovation. Each borrower knows its own θ, but no one else does. Corresponding to each θ is a continuum of borrowers. After it is known how many investment banks have successfully innovated, each borrower approaches either a commercial bank or the capital market for funds; whether all of these borrowers will actually need loans will be determined stochastically only at $t = 1$ according to the probability of project availability for individual borrowers. The random variable that determines the project availability is identically and independently distributed (i.i.d.) across borrowers and independent of θ.

At $t = 1$, each borrower comes to know whether there is a project available, and each borrower's project availability as well as quality (θ) become common knowledge. Thus total loan demand is realized. That is, corresponding to each θ, there is a distinct total loan demand, and across θs the loan demand realizations are i.i.d. random variables. Based on the earlier financing-source choice decisions of borrowers, we now come to know the realized loan demand for commercial banks and the aggregate volume of debt to be underwritten in the capital market. Those who opted to borrow from commercial banks

will be extended monitored loans at an interest factor r_B until the monitoring capacities of the banks they applied to are exhausted; if there is any loan demand left over, it will be satisfied by extending unmonitored loans at an interest factor r_{NB}.[10] We view this descriptively as a process whereby all those seeking loans are viewed as belonging to a homogeneous pool, and the commercial bank selects all at once a random subset of these borrowers to extend monitored loans to at r_B. Thus, prior to the bank's selection of this subset, each borrower views the probability of receiving a monitored loan at r_B as P, with $P \in (0, 1)$ if loan demand exceeds the bank's monitoring capacity, and $P = 1$ otherwise. The interest factors r_B and r_{NB} and the probability P are all derived endogenously in the next section.

Also observed at $t = 1$ is the aggregate order flow D, but not how much of it came from each type of trader. The measure of informed traders, although not directly observable, is inferred. Thus at $t = 1$, each firm chooses its project, the price of each firm's debt is determined, and payoff-enhancing investment decisions by firms are also made (or not). Finally, at $t = 2$ all payoffs are realized and creditors are paid off if possible. Figure 1 summarizes this sequence of events pictorially.

2. Analysis of the Borrowing Firm's Choice of Financing Source

2.1 Cost of borrowing from a commercial bank

If the CB plans to monitor the borrower, then it knows that the borrower will choose the good project w.p. 1, and the CB will be repaid w.p. η. The equilibrium repayment obligation, r_B, thus solves:

$$\eta r_B = 1 + \tau,$$

which yields

$$r_B = \frac{1 + \tau}{\eta}. \tag{1}$$

If the CB does not monitor, then it knows that the probability is θ that the borrower will invest in the good project and $1 - \theta$ that it will invest in the bad project. The equilibrium repayment obligation, rNB, thus solves

$$\theta \eta r_{NB} + \{[1 - \theta] \times 0\} = 1,$$

which yields

$$r_{NB} = 1/\eta\theta. \tag{2}$$

[10] In an ex-post sense, these borrowers would have been better off going to the capital market. However, they are locked into their choice of financing source by this stage.

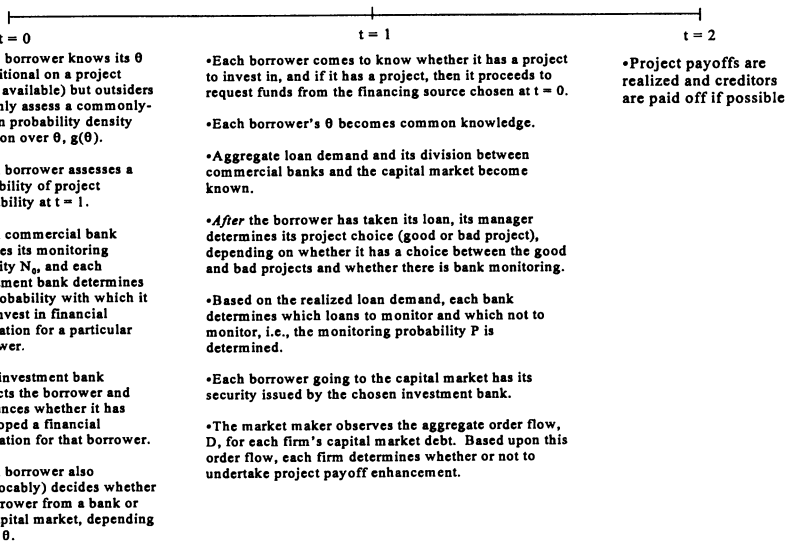

Figure 1
Time line and sequence of events for the model

The borrower's expected payoff if it is monitored is $\eta[Y - r_B]$. Its expected payoff if it is not monitored is $\theta\eta[Y - r_{NB}] + [1 - \theta]R$. To ensure that the borrower prefers to be monitored, we need

$$\eta[Y - r_B] > \theta\eta[Y - r_{NB}] + [1 - \theta]R. \tag{3}$$

We assume that exogenous parameters are restricted such that

$$\eta Y > R + \tau. \tag{PR-1}$$

Given Equation (PR-1), Equation (3) will hold for all $\theta < l - \tau[\eta Y - R]^{-1} \equiv \theta^0$. Note that Equation (PR-1) also guarantees that the borrower would prefer the good project if it could self-finance (which is not possible due to wealth constraints). We also need to ensure that the borrower prefers the bad project with external financing, even when external financing involves the payoff enhancement α. The sufficient condition for this is[11]

$$\eta[Y + \alpha - 1] < R. \tag{PR-2}$$

[11] Equations (PR-1) and (PR-2) imply the joint restriction $\eta Y \in (R + \tau, R + \eta - \eta\alpha)$, which implies $\tau < \eta[1 - \alpha]$.

Next we check to see whether the bank prefers to monitor. If the bank extends a loan at r_B and monitors the borrower, its expected profit will be $1+\tau$. At this loan price, if the bank does not monitor, its expected profit will be $\theta[1+\tau]$, which is less than its expected profit from pursuing a monitoring strategy. Similarly, if the bank extends a loan at r_{NB}, then its expected payoff if it monitors will be $\eta r_{NB} = 1/\theta$, and its expected payoff if it does not monitor will be $\theta\eta r_{NB} = 1 < 1/\theta$. Thus, regardless of the price at which the bank extends a loan, it will strictly prefer to monitor.

Recall that P is the probability that a borrower will receive a loan at r_B from the bank and $1-P$ is the probability that a borrower will receive an unmonitored loan at r_{NB}. Define $\bar{\theta} \equiv [1+\tau]^{-1}$. Then for a borrower with $\theta < \bar{\theta}$, the expected payoff is

$$\begin{aligned}\Pi_B(\theta) &= P\eta[Y - r_B] + [1-P][\theta\eta\{Y - r_{NB}\} + \{1-\theta\}R] \\ &= P\eta Y - P\tau + [1-P][\theta\eta Y + \{1-\theta\}R] - 1. \qquad (4)\end{aligned}$$

2.2 Cost of borrowing in the capital market

Let $\Pr(\phi = f|D)$ denote the conditional probability assessed by the uninformed traders that the informed traders have received a favorable signal. This probability is conditioned on the total demand, D, for the security. A higher realization of D implies a higher probability that the informed traders are in the market, and hence a greater willingness on the borrowing firm's part to engage in the value-enhancing decision. We assume that the value-enhancing decision requires an unverifiable (by outsiders) investment of K. Thus there will be a critical D, call it D^*, such that the firm will make the value-enhancing decision for all $D \geq D^*$ and not otherwise. We will address shortly how D^* is determined. Let $\Pr^*(\phi = f|D)$ represent the probability that the firm invests in value enhancement and actually realizes this enhancement. Thus

$$\Pr^*(\phi = f|D) = \begin{cases} 0 & \text{if } D < D^* \\ \Pr(\phi = f|D) & \text{if } D \geq D^*. \end{cases}$$

We can now write the equilibrium repayment obligation, $\bar{r}_F(D)$, of the firm as a function of the realized demand for its security. It is a solution to

$$\theta\{\Pr^*(\phi = f|D)[\eta\hat{r}_F(D) + \{1-\eta\}\alpha] + [1 - \Pr^*(\phi = f|D)][\eta\hat{r}_F(D)]\} = 1. \qquad (5)$$

Note that in writing Equation (5) we recognize that bondholders get repaid only if the firm is locked into the good project since there is no monitoring in the capital market to deter asset substitution by the firm; the probability of the good project being taken is θ. Moreover, wherever project payoff enhancement occurs, bondholders are repaid

in full in the successful state (this happens w.p. η) and recover $\alpha < 1 < \hat{r}_{\mathrm{F}}(D)$ in the unsuccessful state (this happens w.p. $1-\eta$). Solving Equation (5) yields

$$\hat{r}_{\mathrm{F}}(D) = \frac{1 - \theta \Pr^*(\phi = f|D)[1-\eta]\alpha}{\theta\eta}. \tag{6}$$

Let $\Lambda(D|\Omega)$ represent the cumulative distribution function for D, conditional on the measure of informed traders, Ω. Then the firm's expected payoff is

$$\begin{aligned}\Pi_{\mathrm{F}}(\theta) = \theta\Bigg[\int \eta[Y + \alpha \Pr^*(\phi = f|D) - \hat{r}_{\mathrm{F}}(D)]d\Lambda(D|\Omega) \\ - K\int_{D\geq D^*} d\Lambda(D|\Omega)\Bigg] + [1-\theta]R.\end{aligned} \tag{7}$$

Define $q \equiv q(\Omega) \equiv \int \Pr^*(\phi = f|D)d\Lambda(D|\Omega)$, $r_{\mathrm{F}} \equiv r_{\mathrm{F}}(\Omega) \equiv \int \hat{r}_{\mathrm{F}}(D)d\Lambda(D|\Omega)$, and $\bar{K} \equiv K\int_{D\geq D^*} d\Lambda(D \mid \Omega)$. Then we can write Equation (7) as

$$\Pi_F(\theta) = \theta\eta[Y + \alpha q - r_{\mathrm{F}}] - \theta\bar{K} + [1-\theta]R. \tag{8}$$

We will now determine D^*. Investing K (which happens only when the firm is locked into a good project) when the observed market demand is D produces an expected benefit of $\eta\alpha \Pr(\phi = f|D)$ and has a cost of K. It is transparent that $\Pr(\phi = f|D)$ is increasing in D. Thus there exists a D^* such that

$$\eta\alpha \Pr(\phi = f|D^*) = K, \text{ and}$$

$$\eta\alpha \Pr(\phi = f|D) > K \,\forall\, D > D^*, \text{ and}$$

$$\eta\alpha \Pr(\phi = f|D) < K \,\forall\, D < D^*.$$

2.3 Firm's choice of financing source

The firm will make its financing source choice by comparing $\Pi_B(\theta)$ in Equation (4) with $\Pi_{\mathrm{F}}(\theta)$ in Equation (8). Making this comparison, we obtain the following result.

Proposition 1. *Define*

$$\hat{\theta} \equiv \frac{P[\eta Y - R - \tau]}{\alpha q + P[\eta Y - R] - \bar{K}} \in (0, 1). \tag{9}$$

Then the firm prefers bank financing if its $\theta \leq \hat{\theta}$ and capital market financing if $\theta > \hat{\theta}$. Moreover, all bank-financed borrowers pay a lower interest rate on monitored loans than on nonmonitored loans.

Thus we see that the borrower's choice depends on the publicly observable quality parameter θ. A higher θ means a lower likelihood that the borrower will substitute projects to the lender's detriment, so that θ can be viewed as a representation of the severity of moral hazard. The more severe the moral hazard, the more valuable is the CB's monitoring service. As θ increases, the monitoring becomes less valuable, and at some point the value lost due to not monitoring is more than offset by the expected project payoff enhancement due to capital market financing. At this point the borrower, who has sufficiently high quality, will switch to capital market financing.

3. Analysis of the Decisions of Commercial and Investment Banks

3.1 The commercial bank's choice of lending capacity in a functionally separated banking system

We assume that total loan demand N for each θ is uniformly distributed over $(\underline{N}, \bar{N})$ and that the θ faced by a given CB is uniformly distributed over (0, 1). Thus a CB's *realized* loan demand is N if $\theta \leq \hat{\theta}$ and 0 if $\theta > \hat{\theta}$.[12] Then the CB's choice of lending capacity, N_0, is made to maximize

$$W(\hat{\theta}) = \int_0^{\hat{\theta}} \left[\int_{\underline{N}}^{N_0} \frac{N\tau}{[\bar{N} - \underline{N}]} dN + \int_{N_0}^{\bar{N}} \frac{N_0 \tau}{[\bar{N} - \underline{N}]} dN \right] d\theta - CN_0. \quad (10)$$

There are a few points worth noting about Equation (10). First, the CB's lending to the unmonitored borrowers does not appear here because the CB's expected profit on those loans is zero and hence leaves its overall expected profit W unchanged. Second, CB's expected profit depends both on the realized θ and the realized N. If the CB's monitoring capacity $N_0 > N$ and $\theta \leq \hat{\theta}$, then lending equals demand and some monitoring capacity is wasted. On the other hand, if $N_0 < N$ and $\theta \leq \hat{\theta}$, then lending equals the monitoring capacity and some loans are extended without monitoring. If $\theta > \hat{\theta}$, then lending is zero

[12] Descriptively, one should view this as a market with a finite number of banks, with each bank being uncertain about both the θ of its borrower pool and the loan demand from this pool at the time it determines its lending capacity N. Each bank views the θ it will face as being drawn from a uniform distribution over [0, 1] and the N it will face as being drawn from a uniform distribution over $[\underline{N}, \bar{N}]$. Since there is a finite number of banks, there is only a finite number of relevant θ realizations. Clearly the distribution of N depends on the measure of borrowers associated with each θ as well as the probability distribution of the random variable that determines whether a borrower will have a project to invest in. Since the distribution of N is independent of θ, the simplest case is one in which the measure of borrowers associated with each θ is the same across all θs and the project-availability random variable is i.i.d. across θs.

regardless of N, and all monitoring capacity is wasted. The following proposition follows readily from Equation (10).

Proposition 2. *The CB's optimal monitoring capacity is given by*

$$N_0^* = \bar{N} - C[\bar{N} - \underline{N}][\tau\hat{\theta}]^{-1} < \bar{N}. \tag{11}$$

The probability that a borrower with $\theta \leq \hat{\theta}$ will receive a monitored loan with a repayment obligation of r_B is

$$P = 1 + \frac{\bar{N}}{[\bar{N} - \underline{N}]} \ln\left(\frac{\bar{N}\tau\hat{\theta}}{\bar{N}\tau\hat{\theta} - C[\bar{N} - \underline{N}]}\right) - \frac{C}{\tau\hat{\theta}}\left\{1 + \ln\left(\frac{\bar{N}\tau\hat{\theta}}{\bar{N}\tau\hat{\theta} - C[\bar{N} - \underline{N}]}\right)\right\}. \tag{12}$$

Next we present a corollary that provides some useful comparative statics.

Corollary 1. $\partial N_0^*/\partial\tau > 0$ *and* $\partial P/\partial\tau > 0$.

It is intuitive that N_0^* and P are increasing in τ. Since the CB earns a rent τ on its lending only if it extends a monitored loan, the higher this rent the greater is the investment the CB makes in monitoring. And the greater this investment in monitoring capacity, the higher is the probability that a borrower will receive a monitored loan.

Holding fixed Ω, the measure of informed traders in the capital market, Equations (9), (11), and (12) completely characterize the commercial banking equilibrium with functionally separated banking. Next we turn to the IB's problem.

3.2 The investment bank's problem in a functionally separated banking system

Inspection of Equation (6) reveals that the reduction in the firm's cost of borrowing due to informed trading is captured in the term $\theta \Pr^*(\phi = f|D)[1 - \eta]\alpha/\theta\eta$. The expected value of this is $q(\Omega)[1 - \eta]\alpha/\eta$, where the expectation is taken with respect to D [see the definition following Equation (7)]. This is the cost saving available to the firm with the *existing* security. We assume that the role of financial innovation is to alter security design and *increase* the measure of informed traders from Ω to $\Omega^* > \Omega$. Boot and Thakor (1993) explain how altered security design can achieve this by making more information-sensitive securities available to wealth-constrained informed traders. Define $\Delta \equiv q(\Omega^*) - q(\Omega)$. Then the cost reduction attributable to the financial innovation is $\Delta[1 - \eta]\alpha/\eta$ and its expected value is $\theta\Delta[1 - \eta]\alpha$ for a type-θ borrower, with this expectation taken with

respect to whether the firm will be locked into a good project and whether that project will succeed. In addition to this cost saving, there is an expected enhancement in the firm's project payoff due to the innovation, which is $\theta\eta\Delta\alpha - \theta\bar{K}_0$, where $\bar{K}_0 \equiv \bar{K}(\Omega^*) - \bar{K}(\Omega)$ and $\bar{K}(\Omega^*) \equiv K\int_{D\geq D^*} d\Lambda(D|\Omega^*)$, $\bar{K}(\Omega) \equiv K\int_{D\geq D^*} d\Lambda(D|\Omega)$. If the IB responsible for the innovation is the only one to bring it to market, then it captures all of the borrowing firm's cost saving and payoff enhancement due to the innovation. Thus its reward for the innovation is an increase in its fee revenue by an amount $F \equiv \theta\Delta[1-\eta]\alpha+\theta\eta\Delta\alpha-\theta\bar{K}_0 = \theta[\Delta\alpha - \bar{K}_0] > \xi$, where $\xi > 0$ is what the IB must invest in order to come up with the financial innovation. It is assumed that there are many IBs in the market and any can avail of the financial innovation by investing ξ. For later use, define $S \equiv \Delta\alpha - \bar{K}_0$, so that $F \equiv \theta S$.

Proposition 3. *There does not exist a symmetric pure strategy Nash equilibrium in the game in which multiple IBs compete to innovate. If there are $M > 1$ IBs competing, then the probability, z_s, with which each IB innovates in a mixed-strategy Nash equilibrium with functionally separated banking is*

$$z_s = 1 - \sqrt[M-1]{\frac{2\xi}{F[\underline{N} + \bar{N}][1 - \hat{\theta}]}}. \tag{13}$$

A remark about the interpretation of this equilibrium is in order. Since this is a mixed-strategy equilibrium, each IB is indifferent (in equilibrium) between innovating and not innovating. However, this indifference is based on the IB's assumption that every other IB will innovate with probability z_s. Since the particular IB will then be indifferent, we are free to choose any probability of innovation for it, and the only one that is consistent with the Nash equilibrium assumptions of the other IBs is z_s. We therefore interpret z_s as the probability with which each IB will invest in innovation. For simplicity, we will assume henceforth that $M = 2$.

3.3 The universal bank's problem

With universal banking, the CB and the IB are part of the same bank. Assume that there are two universal banks. Thus the universal bank maximizes the sum of its expected profits from commercial and investment banking [see Equation (10) and Equation (A-2) in the Appendix]. Conditional on the universal bank investing in financial innovation, the total expected profit maximized by the universal bank is

$$zW(\hat{\theta}_2) + [1 - z]W(\hat{\theta}) + \frac{F[\underline{N} + \bar{N}][1 - \hat{\theta}][1 - z]}{2} - \xi \tag{14}$$

where z is the probability with which each universal bank innovates, and $W(\cdot)$, the profit from the commercial bank's lending, was defined in Equation (10). Note that there are two quality cutoffs, $\hat{\theta}$ and $\hat{\theta}_2$. The cutoff $\hat{\theta}$ is that which obtains when only one universal bank innovates, and this is the same as the cutoff without innovation. The reason is that when only one bank innovates, all the benefits of innovation accrue to the bank and the borrower is indifferent between purchasing that innovation and not purchasing it. If both universal banks innovate, then the borrower extracts the entire innovation gain F, and a new cutoff $\hat{\theta}_2$ emerges.

The borrower's expected utility from financial market financing with only one universal bank innovating is $\Pi_{\mathrm{F}}^1(\theta) \equiv \Pi_{\mathrm{F}}(\theta)$, where $\Pi_{\mathrm{F}}(\theta)$ is given by Equation (8), so that the quality cutoff remains $\hat{\theta}$.

The borrower's expected utility from financial market financing with two universal banks present is

$$\Pi_{\mathrm{F}}^2(\theta) = \Pi_{\mathrm{F}}^1(\theta) + z^2\theta S. \tag{15}$$

In writing Equation (15) we recognize that the borrower benefits from the financial innovation only when *both* universal banks innovate (the probability of which is z^2). Now, $\hat{\theta}_2$ is obtained by equating Equations (4) and (15). This yields

$$\hat{\theta}_2 = \frac{P[\eta Y - R - \tau]}{\alpha q + P[\eta Y - R] - \bar{K} + z^2 S}. \tag{16}$$

Observe that $\hat{\theta}_2 < \hat{\theta}$. We now have our main result.

Proposition 4. *The equilibrium probability of financial innovation in a universal banking system, z_u, is lower than the equilibrium probability of financial innovation in a functionally separated banking system, z_s.*

The intuition is as follows. When a functionally separated IB determines whether to innovate, it is unconcerned about the impact the innovation will have on the loan demand faced by a CB. However, when it is the universal bank that determines whether to innovate, it internalizes the depressing effect that the innovation will have on the loan demand faced by its CB unit; this result is independent of the organizational details of the universal bank — whether the IB and the CB are divisions or subsidiaries — and depends only on the fact that the universal bank maximizes the sum of the expected profits of its IB and CB. Consequently the universal bank needs a higher expected profit from the innovation than does a functionally separated IB. Since a positive profit from innovation is available only if the uni-

versal bank in question is the only bank that innovates, the only way to increase the expected profit from innovation is to lower the probability with which each competing bank innovates in a mixed-strategy Nash equilibrium. Note also that it follows from Proposition 4 that the depressing effect of universal banking on financial innovation is dependent on industry structure. With greater concentration in universal banking, z_u is lower since each universal bank internalizes to a greater degree the impact of its own innovation on its commercial banking profits.

Proposition 4 is obtained in a static setting. As we discuss in Section 4A, the propensity of a universal banking system to innovate less is likely to be exacerbated in a dynamic setting.

4. Implications of Analysis

In this section we discuss the implications of our analysis for various aspects of financial system design.

4.1 Intertemporal considerations

An important consideration precluded by our static analysis is reusability of information by CBs. A CB's investment in monitoring is likely to be intertemporally reusable [see Bhattacharya and Thakor (1993) and Boot, Greenbaum and Thakor (1993)]. This means that the cost of monitoring a borrower at date $t + 1$ is likely to be lower than the cost of having monitored the same borrower at date t. The customers of a CB will therefore be more profitable to the CB over time.[13] By contrast, financial innovation yields only a single-shot gain due to imitation by rivals.

When this consideration is introduced in our analysis, we see that a universal bank innovates with an even lower probability since it now imputes a greater cost to the loss in loan demand suffered by its CB due to the financial innovation. Thus intertemporal considerations are likely to strengthen the result that there will be less financial innovation in economies with universal banks.

4.2 Banking scope and capital market development

Perhaps the clearest implication of our analysis is that banking scope — a regulatory choice variable — affects the development of the capital market. In our model this effect arises from the lower incentives

[13] See Greenbaum and Thakor (1995) for empirical evidence. Sharpe (1990) and Rajan (1992) develop models in which the assumulation of proprietary borrower-specific information during the course of a lending relationship creates an informational monopoly for the bank and produces ex post rents.

for financial innovation with universal banking than with functionally separated banking. This stochastic lowering of financial innovation with universal banking means a higher $\hat{\theta}$ and hence fewer borrowers accessing the capital market. With lower aggregate trading volume as well as less financial innovation, we should expect capital markets in economies with universal banking to be less developed than those in economies with functionally separated banking. Moreover, since capital market funding becomes more attractive for borrowers over time due to financial innovation, CBs are likely to lose *more* market share to the capital markets over time in functionally separated financial systems than in universal banking economies. These observations are consistent with the higher incidence of financial innovation and the greater intertemporal loss of market share by CBs in the United States relative to the continental European universal banking economies of Germany, Switzerland, and The Netherlands, for example.

Consistent with the predictions of our model, the corporate bond markets in many of these universal banking economies are not well developed.[14] However, we have assumed in our formal model that universal banks operate in fairly liquid bond markets, which implies that our analysis overstates the importance of bond markets in universal banking economies. Note also that the applicability of our analysis is unaffected by whether the financial contract used is debt or equity. Thus we could readily view our analysis as focusing on the borrower's choice between public equity underwritten by the investment banking arm of a universal bank and private equity offered by the commercial banking arm of a universal bank.

We doubt that the architects of the Glass–Steagall Act foresaw the enormously positive impact the act would have on the development of U.S. capital markets or on the incentives for financial innovation. The act had its roots in the desire to limit the power of banks, reduce conflicts of interest, and limit the scope of the deposit insurance safety net. Nonetheless, our analysis provides a framework within which to understand the unintended consequences of banking scope legislation like the Glass–Steagall Act.

4.3 Path dependence in the evolution of financial systems

Financial innovation is likely to be path dependent. Although we have assumed that an investment in financial innovation will succeed with probability one, it would be straightforward to let the success proba-

[14] The Netherlands is particularly illustrative. The government bond market there is very liquid, suggesting that an adequate infrastructure for bond trading exists. However, the corporate bond market is not as well developed, suggesting that the incentives for (private) universal banks to innovate and facilitate development of the corporate bond market may be particularly weak.

bility be less than one. While this by itself will not alter our analysis, we could incorporate the insight offered by Gale (1992) and argue that the success probability will be a function of the sophistication of financial market participants. The ability of market participants to appreciate the payoff implications of a new security will likely depend on their experience with existing securities, the attributes of which may depend on the development of the capital market. Thus the probability of financial innovation will be higher in a better-developed capital market. Over time this will lead to differences in the rate at which financial innovation proceeds in different markets. This difference in the pace of financial innovation further widens the development gap between better-developed capital markets and their less-developed counterparts.

This implies that even if the regulation of banks and capital markets were to be perfectly harmonized internationally, different financial systems are likely to display disparate levels of financial innovation and differing fractions of total credit allocation accounted for by CBs simply due to disparities in the sophistication of their capital markets. Moreover, how sophisticated a capital market is at date t is likely to depend on the history of financial innovation until date t. A financial system that has historically been dominated by universal banks is likely to have a poorer history of financial innovation, according to our earlier arguments. This appears consistent with the different patterns of capital market development in continental Europe and the United States.

4.4 Commercial banking fragmentation implications

Greater fragmentation of commercial banking is typically taken to mean greater competition among CBs. In our model this implies a lower τ for each CB. From Equation (9) we know that $\partial\hat{\theta}/\partial\tau < 0$. Thus increased fragmentation in commercial banking will lead to an increase in $\hat{\theta}$ and hence more business for CBs.[15] This elevates borrower welfare as well as the average quality of bank loans (since $\hat{\theta}$ increases, the average quality of bank loans increases with it).[16] Nevertheless, Corollary 1 also tells us that the probability, P, of extending a monitored loan declines as τ decreases. Moreover, the effect of in-

[15] The observation that more competitive commercial banking results in more use of banks may seem counter to actual experience, not only in universal-banking Germany but also in functionally separated Japan. This is probably because increased bank competition may be accompanied (and perhaps caused) by improved information technology, which also increases the efficiency of public markets.

[16] One may wonder by CBs do not reduce τ for borrowers with θ just above $\hat{\theta}$. The reason they don't do this is that borrowers precommit to a choice of financing source prior to knowing whether they will receive a monitored loan.

creased competition on the bank's investement in monitoring capacity is ambiguous since a higher $\hat{\theta}$ implies a higher N_0^* ceteris paribus but a lower τ (which leads to a higher $\hat{\theta}$) diminishes N_0^* for a fixed $\hat{\theta}$.

4.5 Implications of increased competition in investment banking

Fragmentation and the resulting increased competition in investment banking will diminish the inclination of any IB to introduce a financial innovation. Recall that the probability of each IB innovating is chosen such that the net present value of the innovation to the IB is zero. From Equation (13) we see that $\partial z_s/\partial M < 0$. More importantly, however, the probability that there will be any innovation at all — the probability that at least one out of M IBs will innovate — declines as M increases.[17] Hence, increased competition among IBs leads to stochastically lower innovation.

4.6 Overall financial system design

Our analysis shows that financial systems with universal banking can be expected to innovate less and have capital markets that display lower development than financial systems with functionally separated banking. Since an important role of the financial market in our model is to provide informational feedback to managers of firms that facilitate improved real decisions, borrowers make better real decisions on average in functionally separated financial systems.

On the other hand, there is on average better attenuation of asset-substitution moral hazard in a financial system with universal banking because a larger measure of borrowers use CBs. The welfare implications of financial system design are therefore ambiguous. Of course, because our capital market model has exogenously specified security demand from liquidity traders, it is not amenable to welfare analysis. However, one could adapt the model in such a way that it is the firm rather than the liquidity traders who provide compensation to the informed traders: welfare analysis would then be possible.

Stepping outside our model, a factor that might favor universal banking is related to scope economies based on information sharing made possible by the marriage of commercial and investment banking. However, potential gains from scope economies could be

[17] To see this, note from Equation (13) that Pr (at least one bank innovates) $= 1 - [1 - z_s]^M$, where z_s is defined in Equation (13) and $0 < \frac{2\xi}{F[\bar{N}+\underline{N}][1-\hat{\theta}]} < 1$. It can now be shown that ∂ Pr (at least one bank innovates)$/\partial M < 0$. In a takeover bidding context, Spatt (1989) proves a similar claim. See also Thakor (1996) for a proof in a credit rationing context.

vitiated by conflicts of interest in a universal bank [see Kroszner and Rajan (1994b) and Rajan (1993)].

4.7 Mixed financial systems

We have considered functionally separated banking and universal banking as two extremes. What about "mixed" financial systems in which stand-alone IBs and CBs compete with universal banks?

We believe that stand-alone banks would be competitively disadvantaged in a universal banking system for two reasons. First, scope economies would give universal banks a competitive edge over their stand-alone counterparts. In the context of our model, one way to introduce scope economies would be to assume that if there is any redundant monitoring capacity in the CB unit of the universal bank, it could be used to support the underwriting activities of the IB. This would lower expected underwriting costs, and some of the savings could be passed along to the universal bank's customers.

Second, although we don't have a good theory that provides testable links between organization size and its influence-peddling ability, casual observation suggests that a larger universal bank typically deals with larger, more politically visible clients, has "deeper pockets" and enjoys greater implicit "too big to fail" (TBTF) protection than its smaller (particularly stand-alone) competitors [see Kane (1996)]. Thus it seems reasonable to posit that large universal banks have greater influence over regulators than (smaller) stand-alone IBs or CBs, which means that regulatory policy could also be slanted in favor of universal banks.[18] For example, financial innovations where scope economies could be exploited more fully may be favored over others when it comes to regulatory approval. A good example is commercial paper with backup loan commitments. The universal bank can underwrite the commercial paper issue and also sell the backup loan commitment.

Both of these considerations imply that stand-alone banks, even though viable, are unlikely to be major players in universal banking economies, an observation that appears consistent with what we observe. Hence it seems improbable that overall financial innovation in a universal banking system with some stand-alone CBs and IBs could match the financial innovation in a functionally separated financial system.

[18] Of course, if there are numerous small banks that collectively represent a large fraction of banking industry assets, these banks could coordinate to collectively lobby regulators and politicians. This has been the case in the United States. However, unlike the United States, in economies traditionally dominated by large universal banks, small stand-alone banks are unlikely to represent a sufficiently large fraction of industry assets to have significant political clout. And they would be further hampered by the fact that the interests of stand-alone CBs would differ from those of stand-alone IBs, leading to collective lobbying being frustrated by the usual coordination problems.

4.8 Global competition among institutions from different systems

At a more general level, the issue of stand-alone investment banks competing with universal banks raises the important issue of competitiveness of different financial systems in an increasingly integrated global economy. While cross-border competition is limited at present, it does exist nonetheless. How would a bank-dominated (universal banking) system compete with a market-dominated (functionally separated) system? This is an interesting question for future research.

Our analysis does suggest, however, that IBs from functionally separated financial systems would have an innovation-based advantage in competing with universal banks from universal banking systems. Thus if these IBs were allowed to operate in universal banking economies, they would wrest some market share away from local universal banks, particularly when it came to large corporate borrowers seeking capital market funding. Interestingly this is precisely what has happened with U.S. banks entering Germany. We quote from the *Wall Street Journal* (October 1995):

> Bavaria's $4 billion sale of its electric utility was the largest deal in Germany last year. But the Bavarian state government didn't choose Deutsche Bank AG or another big German bank to lead the auction. Instead, it turned to Lehman Brothers, Inc.
>
> The choice of a American adviser for such a significant deal shows the growing influence of U.S. investment banks in Germany.
>
> The breakthrough came with German reunification in 1990, when the nation's financing needs soared. U.S. banks stood ready to fill that need. And since then, whether it has been privatizations, mergers or stock offerings, they have won one assignment after another. They have come to dominate futures and options trading on the Frankfurt Stock Exchange. And they have scored points for introducing aggressive American concepts such as structured notes and leveraged buyouts to the staid German market.
>
> It's that innovative edge, U.S. bankers say, that sets them apart.

5. Conclusion

We have focused on the financial innovation implications of financial system design. Our main findings and observations are summarized below.

There is an observable quality cutoff such that borrowers with observable qualities below that are funded by commercial banks and borrowers with observable qualities above that are funded in the cap-

ital market. As commercial banking becomes more competitive, this cutoff increases.

There exists a mixed-strategy Nash equilibrium in the financial innovation game such that each competing investment bank invests in financial innovation with some probability less than one. The equilibrium probability of innovation is lower in a financial system with universal banking than in a financial system with functionally separated banking.

The evolution of a financial system is likely to be path-dependent. Well-developed financial systems provide stronger incentives for financial innovation and develop faster.

Banks are likely to lose more market share over time to capital markets in financial systems with functionally separated banking than in a universal banking system.

The choice of financial system design rests on the trade-off between the superior attenuation of asset-substitution moral hazard in a universal banking system versus superior financial innovation and better real decisions in a functionally separated financial system.

Perhaps the most significant point of our article is that there is a vital link between the behavior of commercial banks and developments in capital markets, and that any discussion of financial system design must adopt an essentially integrated approach.[19] Moreover, bank regulation and capital market regulation, which are typically the responsibilities of different regulatory agencies, should be conducted in an integrated manner.

Future research should perhaps attempt to join together the implications of financial system design derived in recent articles. For example, Allen and Gale (1995) conclude that bank-dominated financial systems provide better intergenerational risk sharing and market-dominated systems provide better cross-sectional risk sharing. That is, generally speaking, financial innovation should be thought of more broadly as improving risk sharing and providing tax advantages, in addition to increasing information sensitivity. It would be interesting to incorporate risk-sharing considerations in the approach we have taken. In particular, including liquidity demand considerations [e.g., as in Kahn and Winton (1996)] in a framework like ours could open the door for interesting welfare analyses. Moreover, one could also consider innovations that have *synergies* with bank lending, such as swaps and forward contracts. Our analysis suggest that commercial

[19] Benveniste, Singh and Wilhelm (1993) provide interesting empirical support for the notion that commercial banks and capital markets are significantly linked. They document that the failure of Drexel, Burnham and Lambert led to an abnormal increase in the prices of banks whose loans were viewed as close substitutes for the junk bonds underwritten by Drexel.

banks would be aggressive in introducing them, which is what we observe.

Appendix

Proof of Proposition 1. We know that $\hat{\theta}$ solves $\Pi_B(\theta) = \Pi_F(\theta)$, where $\Pi_B(\theta)$ is defined in Equation (4) and $\Pi_F(\theta)$ in Equation (8). Thus the borrower prefers capital market funding to a CB loan if

$$\theta\eta[Y+\alpha q-r_{\mathrm{F}}]-\theta\bar{K}+[1-\theta]R > P\eta Y-P\tau+[1-P][\theta\eta Y+\{1-\theta\}R]-1.$$

Substituting $r_{\mathrm{F}} = \frac{1-\theta q[1-\eta]\alpha}{\theta\eta}$ in the above inequality and performing a few algebraic manipulations, we obtain the result that the borrower strictly prefers capital-market funding if

$$\theta > \frac{P[\eta Y - R - \tau]}{\alpha q + P[\eta Y - R] - \bar{K}} \equiv \hat{\theta},$$

prefers CB financing if $\theta < \hat{\theta}$, and is indifferent if $\theta = \hat{\theta}$. Moreover, it is straightforward to show that $\hat{\theta} < \theta^0$. Now, define $\bar{\theta} \equiv [1+\tau]^{-1}$. Then it is transparent that $r_{\mathrm{B}} < r_{\mathrm{NB}} \forall \theta < \bar{\theta}$. It is easy to show that $\hat{\theta} < \bar{\theta}$. Thus all those who apply for bank loans find that $r_{\mathrm{B}} < r_{\mathrm{NB}}$. This completes the proof. ■

Proof of Proposition 2. Performing the necessary integration, Equation (10) can be written as

$$W(\hat{\theta}) = \frac{\hat{\theta}\tau[N_0^2 - \underline{N}^2]}{2[\bar{N} - \underline{N}]} + \frac{\hat{\theta}\tau N_0[\bar{N} - N_0]}{[\bar{N} - \underline{N}]} - N_0 C.$$

The first-order condition, $\partial W(\hat{\theta})/\partial N_0 = 0$, yields

$$\frac{\hat{\theta}\tau\bar{N}}{[\bar{N} - \underline{N}]} - \frac{\hat{\theta}\tau N_0^*}{[\bar{N} - \underline{N}]} - C = 0,$$

which then gives us Equation (11). The second-order condition is

$$\partial^2 W(\hat{\theta})/\partial N_0^2 = \frac{-\tau\hat{\theta}}{[\bar{N} - \underline{N}]} < 0.$$

It is transparent from Equation (11) that $N_0^* < \bar{N}$.

To derive P, note that

$$P = \Pr(\text{no shortage of monitoring capacity}) \times \Pr(\text{each loan will be monitored when there is no monitoring capacity shortage})$$

$$
\begin{aligned}
&+ \Pr(\text{shortage of monitoring capacity}) \times \Pr(\text{loan will be monitored when there is a capacity shortage}) \\
&= \Pr(\text{no monitoring capacity shortage}) \times 1 \\
&\quad + \Pr(\text{monitoring capacity shortage}) \times \frac{\text{monitoring capacity}}{\text{loan demand}} \\
&= \int_{\underline{N}}^{N_0^*} \frac{1}{[\bar{N} - \underline{N}]} dN + \int_{N_0^*}^{\bar{N}} \frac{N_0^*}{N[\bar{N} - \underline{N}]} dN \\
&= \frac{[N_0^* - \underline{N}]}{[\bar{N} - \underline{N}]} + \frac{N_0^* \ln(\bar{N}/N_0^*)}{[\bar{N} - \underline{N}]}. \qquad \text{(A1)}
\end{aligned}
$$

Substituting for N_0^* from Equation (12) into Equation (A1) yields Equation (12). ■

Proof of Corollary 1. Differentiating Equation (12) with respect to τ gives

$$\partial N_0^* / \partial \tau = \frac{C[\bar{N} - \underline{N}]}{\tau^2 \hat{\theta}} > 0.$$

Moreover, differentiating Equation (12) with respect to τ and doing a little algebra shows that

$$\partial P / \partial \tau > 0.$$ ■

Proof of Proposition 3. To show that there cannot be a symmetric pure strategy Nash equilibrium in the innovation game, suppose that we conjecture that no IB innovates in equilibrium. Then it must pay for one IB to innovate since it will have a monopoly on the innovation and therefore earn positive expected profit. Thus no innovation cannot be an equilibrium. Next, suppose that it is an equilibrium for each IB to innovate w.p. 1. Then no IB can profit from the innovation and hence cannot recover its investment of ξ in innovation. Thus it cannot be an equilibrium for each IB to innovate w.p. 1.

Let $z \in (0, 1)$ be the probability with which each IB innovates in a symmetric mixed strategy Nash equilibrium. Consider a particular IB. Its expected profit from innovation is

$$[1 - \hat{\theta}] \times \frac{[\bar{N} + \underline{N}]}{2} \times [1 - z]^{M-1} \times F - \xi. \qquad \text{(A2)}$$

Note also that the quality cutoff $\hat{\theta}$ in (A2) is the same as the quality cutoff without financial innovation. The reason is that Equation (A2) is relevant only for the case in which only one IB innovates. As indicated in Section 2.4, in that case, all the benefits of innovation accrue to

the investment bank and the quality cutoff $\hat{\theta}$ remains unaffected. In writing Equation (A2), note that the expected credit demand faced by the IB is

$$\int_{\hat{\theta}}^{1} \int_{\underline{N}}^{\bar{N}} \frac{N}{[\bar{N} - \underline{N}]} dN \, d\theta = \frac{[1 - \hat{\theta}][\bar{N} + \underline{N}]}{2}.$$

Moreover, the IB in question can profit from its financial innovation only if no other IB innovates. Since the probability that an IB will not innovate is $[1 - z]$ and there are $M - 1$ other IBs, the probability that the remaining $M - 1$ banks will not innovate is $[1 - z]^{M-1}$.

To obtain a symmetric mixed strategy Nash equilibrium, we have to ensure that the IB is indifferent between innovating and not innovating. Since the IB's expected profit from not innovating is zero, this means the required equilibrium condition is

$$\frac{[1 - \hat{\theta}][\bar{N} + \underline{N}][1 - z]^{M-1} \times F}{2} - \xi = 0. \tag{A3}$$

Solving Equation (A3) yields Equation (13). ■

Proof of Proposition 4. With universal banking, the bank's objective is to maximize Equation (14). Let z_u be the probability with which each universal bank innovates in a mixed strategy Nash equilibrium. Note that the rule by which innovation rents are shared between the IB and the borrower is immaterial to the analysis. From our earlier analysis, $\hat{\theta} > \hat{\theta}_2$.

Consider now a particular universal bank and assume that there are two universal banks in the market. If the universal bank in question innovates, its expected profit is

$$z_u[W(\hat{\theta}_2) + 0] + [1 - z_u]\left[W(\hat{\theta}) + \frac{F\{\bar{N} + \underline{N}\}\{1 - \hat{\theta}\}}{2}\right] - \xi. \tag{A4}$$

Note that the probability that the other universal bank will innovate is z_u, and in this case each bank earns zero profits in investment banking and an expected profit of $W(\hat{\theta}_2)$ from commercial banking. The probability that the other universal bank will not innovate is $1 - z_u$, and in this case the universal bank in question earns an expected profit of $\frac{F\{\bar{N}+\underline{N}\}\{1-\hat{\theta}\}}{2}$ on its innovation and an expected profit of $W(\hat{\theta}_1)$ on its CB lending. If the universal bank in question does not innovate, then its expected profit is

$$z_u[W(\hat{\theta})] + [1 - z_u][W(\hat{\theta})]. \tag{A5}$$

A key difference between Equation (A4) and Equation (A5) is that now

if the other universal bank does not innovate, then no bank innovates and the expected profit on commercial bank lending is $W(\hat{\theta})$ since the quality cutoff is $\hat{\theta}$.

Now, z_u is obtained by setting Equation (A4) equal to Equation (A5). Solving this gives us

$$z_u = \frac{FN_m[1-\hat{\theta}] - \xi}{FN_m[1-\hat{\theta}] + 2W(\hat{\theta}) - W(\hat{\theta}) - W(\hat{\theta}_2)} \tag{A6}$$

where $N_m \equiv \frac{[\bar{N}+\underline{N}]}{2}$. Note that $W(\hat{\theta}) < W(\hat{\theta}_2)$ since W is increasing in $\hat{\theta}$. We wish to compare Equation (A6) and Equation (13). Note first that Equation (13) can be stated as

$$z_s = \frac{FN_m[1-\hat{\theta}] - \xi}{FN_m[1-\hat{\theta}]}. \tag{A7}$$

Comparing Equation (A6) and Equation (A7) and recalling that $W(\hat{\theta}) > W(\hat{\theta}_2)$, we see that $z_u < z_s$. ■

References

Allen, F., 1990, "The Market for Information and the Origin of Financial Intermediation," *Journal of Financial Intermediation*, 1, 3–30.

Allen, F., 1992, "Stock Market and Resource Allocation" in C. Mayer and X. Vives (eds.), *Capital Markets and Financial Intermediation*, Cambridge University Press, Cambridge.

Allen, F., and D. Gale, 1988, "Optimal Security Design," *Review of Financial Studies*, 1, 229–263.

Allen, F., and D. Gale, 1991, "Arbitrage, Short Sales, and Financial Innovation," *Econometrica*, 59, 1041–1068.

Allen, F., and D. Gale, 1994, *Financial Innovation and Risk Sharing*, MIT Press, Cambridge, Mass.

Allen, F., and D. Gale, 1995, "A Welfare Comparison of the German and U.S. Financial Systems," *European Economic Review*, 39, 179–209.

Benveniste, L. M., M. Singh, and W. J. Wilhelm, Jr., 1993, "The Failure of Drexel Burnham Lambert: Evidence on the Implications for Commercial Banks," *Journal of Financial Intermediation*, 3, 104–137.

Berlin, M., K. John, and A. Saunders, 1994, "Bank Equity Stakes in Borrowing Firms and Financial Distress," working paper, Indiana University.

Besanko, D., and A. V. Thakor, 1992, "Relationship Banking, Deposit Insurance, and Portfolio Choice," in C. Mayer and X. Vives (eds.), *Capital Markets and Financial Intermediation*, Cambridge University Press, Cambridge.

Bhattacharya, S., and V. Nanda, 1996, "Client Discretion, Switching Costs and the Introduction of New Financial Products," working paper, University of Michigan.

Bhattacharya, S., and G. Chiesa, 1995, "Proprietary Information, Financial Intermediation, and Research Incentives," *Journal of Financial Intermediation*, 4, 328–357.

Bhattacharya, S., and A. V. Thakor, 1993, "Contemporary Banking Theory," *Journal of Financial Intermediation*, 3, 2–50.

Boot, A. W., and A. V. Thakor, 1993, "Security Design," *Journal of Finance*, 48, 1349–1378.

Boot, A. W., and A. V. Thakor, 1997, "Financial System Architecture," *Review of Financial Studies*, 10, 693–733.

Boot, A. W., S. I. Greenbaum, and A. V. Thakor, 1993, "Reputation and Discretion in Financial Contracting," *American Economic Review*, 83, 1165–1183.

Dinc, S., 1994, "Integration of Financial Systems and Institutional Path Dependence," working paper, Stanford University.

Duffie, D., and R. Rahi, 1995, "Financial Market Innovation and Security Design," *Journal of Economic Theory*, 65, 1–42.

The Economist, 1994, "Central Europe's Model Bank," August 27–September 2.

Fishman, M., and K. Hagerty, 1995, "The Incentive to Sell Financial Market Information," *Journal of Financial Intermediation*, 4, 95–115.

Gale, D., 1992, "Information Capacity and Financial Collapse," in C. Mayer and X. Vives (eds.), *Capital Markets and Financial Intermediation*, Cambridge University Press, Cambridge.

Greenbaum, S. I., and A. V. Thakor, 1995, *Contemporary Financial Intermediation*, Dryden Press, Fort Worth, Texas.

Holmstrom, B., and J. Tirole, 1993, "Market Liquidity and Performance Monitoring," *Journal of Political Economy*, 101, 678–709.

Kanatas, G., and J. Qi, 1994, "Should Commercial Banks be Prohibited from Underwriting Corporate Securities?," working paper, University of South Florida.

Kahn, C., and A. Winton, 1996, "Ownership Structure, Liquidity Demand, and Shareholder Monitoring," working paper, University of Illinois and Northwestern University.

Kane, E. J., 1996, "Foundations of Financial Regulation," paper presented at a conference, *Monetary and Financial Integration in an Expanding (N)AFA: Organization and Consequences*, Toronto, Canada.

Kreps, D. M., and J. A. Scheinkman, 1983, "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics*, 14, 326–337.

Kroszner, R. S., and R. G. Rajan, 1994a, "Is the Glass-Steagall Act Justified? A Study of the U.S. Experience with Universal Banking Before 1933," *American Economic Review*, 84, 810–832.

Kroszner, R. S., and R. G. Rajan, 1994b, "The Role of Firewalls in Universal Banks," working paper, University of Chicago.

Nanda, V., and Y. Yun, 1994, "Competititon and Cooperation Among Investment Banks in the Market for Financial Innovations," working paper, University of Southern California.

Neave, E. H., and L. D. Johnson, 1993, "Governance and Financial System Organization," Working Paper 93-26, School of Business Research Program, Queen's University.

O'Hara, M., 1993, "Real Bills Revisited: Market Value Accounting and Loan Maturity," *Journal of Financial Intermediation*, 3, 51–76.

Pound, J., 1997, "The Resurrection of NewTech," excerpted from *The New Public Corporation*, forthcoming.

Puri, M., 1994, "The Long-Term Default Performance of Bank Underwritten Security Issues," *Journal of Banking and Finance*, 18, 397–418.

Rajan, R. G., 1992, "Insiders and Outsiders: The Choice Between Informed and Arm's-Length Debt," *Journal of Finance*, 47, 1367–1423.

Rajan, R. G., 1993, "A Theory of the Costs and Benefits of Universal Banking," working paper, University of Chicago.

Riddiough, T., 1997, "Optimal Design and Governance of Asset-Backed Securities," *Journal of Financial Intermediation*, 6, 121–152.

Sabani, L., 1992, "Market Oriented Versus Bank Oriented Financial Systems: Incomplete Contracts and Long Term Commitments," unpublished manuscript, Universita' di Roma "La Sapienza" Dipartimento di Economia Pubblica and Trinity College, Cambridge.

Sharpe, S. A., 1990, "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships," *Journal of Finance*, 55, 1069–1087.

Spatt, C. S., 1989, "Strategies Analyses of Takeover Bids," in S. Bhattacharya and G. Constantinides (eds.), *Financial Markets and Incomplete Information: Frontiers of Modern Financial Theory, Volume 2*, Rowman & Littlefield Publishers, Inc., N.J.

Thakor, A. V., 1996, "Capital Requirements, Monetary Policy and Aggregate Bank Lending: Theory and Empirical Evidence," *Journal of Finance*, 51, 279–324.

Titman, S., and A. Subrahmanyam, 1996, "Information, Resource Allocation and the Development of Financial Markets," unpublished manuscript, Boston College and UCLA.

Yosha, O., 1995, "Information Disclosure Costs and the Choice of Financing Source," *Journal of Financial Intermediation*, 4, 3–20.