What Are Financial Intermediaries Paid For?

The old story of the life of a banker was simple: borrow at 3%, lend out at 5%, be at a golf course at four. Far from being just a joke, this story is a representation of the mission of financial intermediation. A representation that never was, but it has an appealing simplicity. The question one needs to ask is: What did the banker do to earn the 2% spread?

One of my former students, now a bond portfolio manager, approached me with a question that represented the other side of the coin. She learned about mortgage-backed securities and the dangers of negative convexity. Her question was: Why would anyone purchase a security that offered no appreciation potential with lower interest rates and a threat of loss with higher interest rates? Why then does a more than a trillion-dollar mortgage-backed securities market exist?

Anthony Saunders (1994, Chap. 3) discusses the fundamental question concerning the nature of the financial industry: ''Why are financial intermediaries special?'' Undoubtedly, financial intermediaries—that is, banks, insurance companies, securities firms, investment banks, and mutual funds—are singled out for special regulatory attention. Saunders acknowledges this situation and wonders what makes the industry so different that such regulatory attention is necessary.

Here is a summary of his findings. In a world without financial intermediaries, the national economy's flow of funds would concentrate on the exchange between the household sector, which contains net savers, and the production sector (corporations and other businesses), which contains net borrowers. However, because of monitoring costs, liquidity costs, price risk, and similar financial reasons, the average household saver may view investment in corporate securities as unattractive. Saunders (1994) provides empirical evidence of such aversion, and Ostaszewski (1995) discusses the topic as well. For its part, the production sector, especially if it succeeds in growing its business, is in constant need of funds for financing such growth. Thus, financial intermediaries appear to ''grease the wheels of commerce.'' They provide securities, which are indeed needed by the household sector (thus assuming a short position in those securities), and use the funds to purchase securities supplied by the corporate sector (therefore assuming a long position in those securities). There are also other types of situations where this short/long portfolio is automatically created by intermediation; for example, in payment facilitation, where the intermediary is short accounts receivable and long accounts payable. We claim that creation of such a short/long portfolio is the essence of the intermediation business. The result is that the efficiency of use of capital in the economy is improved, and economic growth is facilitated. Clearly, protection of the resulting link between the savers and producers is one of the major missions of financial intermediaries regulation.

Traditionally, the resulting function of financial intermediaries is described as the ''spread business.'' However, this description addresses only the manner in which banks, insurance companies, or investment companies are paid. The nature of their business lies in assuming a long position in corporate securities (and similar *primary securities,* as Saunders terms them), and a short position in securities issued to the household sector (i.e., bank accounts, insurance policies, investment accounts, and so forth). What we propose here is somewhat of a change in perspective. The short/long position can, and should, be viewed as a *derivative security.* This claim was already made, in a limited form, in an earlier work (Ostaszewski 1995).

The perception of our business as a ''spread business'' has some hidden traps. When Michael Milken nearly single-handedly created the original issue junk

bond boom in the 1980s, it was based on the thesis that junk bonds provided a higher level of return than did Treasuries or investment-grade bonds, with lower levels of risk. Return was measured by the expected return, and risk by the standard deviation of return. So junk bonds looked like a ''dominant'' security with respect to Treasuries: more return with less risk. The logical extension of that idea is what we could call *the Milken arbitrage*: If junk dominates investment grade bonds, can we sell AAA bonds short and use the proceeds to buy junk? This can't be done in the market, but there's nothing to stop an insurance company (such as Executive Life) from trying it. As soon as a company acquires the highest credit rating, it can sell deferred annuities that (implicitly) have that rating, while investing the proceeds in junk bonds—thereby earning a comfortable spread. Is there any risk? The short position is riskier than the long position, so the insurer looks very comfortable.

But let us look at some details lurking beneath the surface. The return of the short/long position is uncertain, and should be modeled as such. The most direct approach is to view it as a random variable. The variance of the combined position's return isn't the difference of variances; it is the sum of the two variances *minus* twice the covariance of the two returns. Long Treasury bonds, or long AAA corporate bonds, are not positively correlated with below-investment-grade bonds. (In fact, it is possible to make a case for a negative correlation.) Therefore, the combined short/long position could end up being more than twice as risky as the individual pieces (if variance is used as a measure of risk).

This ''pseudo-arbitrage'' is actually a derivative security. The short AAA/long junk position is additionally enhanced by numerous options offered to holders of deferred annuities. The company is paid for those options: granting them to the customers allows the company to credit less than a regular corporate AAA bond pays, that is, earn the spread in this ''spread business.'' But just as there is no free lunch, there is no free spread margin. One must accept the increased risk in the form of large variance of returns in order to get the spread.

Financial intermediaries are not in the spread business. We are in a derivative securities business. Our job is to create securities that will match the supply of savings with the demand for securities. And these can, and should, be termed derivatives. Derivative securities have received a significant amount of adverse publicity recently, but there is no escaping them—they are an integral part of financial life. The concept of a derivative traditionally includes options, futures, forwards, swaps, and similar exotic securities. Hull (1993) provides an excellent overview of modern techniques of valuation of exotic derivatives. However, those techniques invariably date back to the seminal analysis of Black and Scholes (1973), which included not just the famed formula for the price of a European call, but the powerful idea that valuation of financial instruments must have provision for contingent claims.

Let us recall some derivatives in the classical meaning of the word. A *call option* is the right to purchase a security at a predetermined price in some period (American option) or a point in time (European option) in the future. A *put option* is a similar right to sell a security. A *forward contract* is a purchase agreement for goods or securities that is signed now, while the actual delivery of the good or securities occurs at a predetermined time in the future. A *future contract* is a traded forward contract that is secured by marking its value to the market value of the underlying good or security to its market price, and by a margin balance. Finally, a *swap* is an agreement to exchange cash flows produced by securities (e.g., a fixed coupon long-term bond and a floating coupon bond) without actual delivery of securities. There are also various mortgage derivatives created out of pools of mortgages (providing rights to specified portions of cash flows produced by such pools of mortgages), or so-called ''structured notes,'' which are custom-structured securities designed to meet the needs of trading parties. Undoubtedly, a structured note has exactly the same intermediation function as instruments issued by banks or insurance companies.

In contrast, modern analysis of financial intermediation as performed by insurance companies shows that very traditional insurance policies contain derivatives such as options or forwards. For example, traditional life insurance policies and deferred annuities provide a minimum interest rate guarantee. If the guarantee is for 4%, this is equivalent to the right of a policyholder to purchase a 4% bond (i.e., a call on a 4% bond). Such right becomes extremely valuable when interest rates drop below 4%, but it does have a market value under any market circumstances. Similarly, if a deferred annuity provides for the credited rate to follow an interest rate index, it gives the policyholder an option to sell lower interest rate bonds at par or near par and purchase bonds paying the index-related rate at par or near par.

Smith (1982) shows that a life insurance policy can be viewed as an option package. Doherty and Garven

(1986) modeled the property-liability insurance as a bundle of long- and short-call options. We will return to such models in the later chapters.

The essential feature of various derivatives securities discussed here is that their cash flows are not deterministic, but uncertain, contingent upon events that trigger them. Generally, the triggering events are a function of some ''underlying'' security, such as a stock or a bond. Therefore, valuation of derivatives is also commonly referred to as valuation of contingent claims (i.e., claims to said cash flows). What this perspective on financial intermediation implies is that valuation of contingent claims must become an integral part of management of financial intermediaries. The proper domain of asset-liability management is therefore the study of the interaction of financial intermediaries' assets and securities issued by them, commonly called their liabilities. It is an emerging area of insurance management, although already established in terms of its significance. ALM is invariably tied to valuation of derivative securities and is the essence of financial intermediation.

Core Functions

Merton and Bodie (1995), as well as Jordan (1996), present a conceptual framework very closely related to what is being proposed here. They propose viewing the financial environment from the *functional perspective.* This implies that functions rather than institutions should be the conceptual ''anchor'' for understanding financial intermediation. The main reasons for adopting such formulae are that (1) it is the function that remains relatively stable over time and place (i.e., similar functions can be identified in various countries, at various times) and (2) the form must adapt to the function eventually. Merton and Bodie (1995) identified six core functions:

- To provide ways of *clearing and settling payments* to facilitate trade.
- To provide a mechanism for the *pooling of resources* and for the subdividing of shares in various enterprises.
- To provide ways to *transfer economic resources* through time, across borders, and between industries.
- To provide ways of *managing risk*.
- To provide *price information* to help coordinate decentralized decision making in various sectors of the economy.
- To provide ways of *dealing with the incentive problems* created when one party to a transaction has information that the other party does not or when one party acts as agent for another.

One important aspect emerging from this analysis is the question of the relationship between the institutional intermediaries and capital markets. Merton and Bodie (1995) view the evolution of this relationship as an innovation spiral in which organized markets and intermediaries compete with each other in a static sense, and complement each other in a dynamic sense. In other words, any instrument issued by intermediaries is compared by their customers to marketable securities available for purchase. At the same time, intermediaries must create their products out of the supply of securities available and, thus, must adjust to the changing nature of that supply. Let us further examine the Merton-Bodie model and, given our perspective, look for the derivatives.

Clearing and Settling Payments

This function exists mainly because delivery of goods or services by the producer creates a financial liability (accounts payable) on the balance sheet of the receiving party, and an asset on the balance sheet of the delivering party (accounts receivable). The derivative created by an intermediary is the short/long position described previously—short accounts payable, long accounts receivable. There is, of course, a greater variety of functions performed here, including netting arrangements, use of collateral, delivery-versus-payment, or extension of credit. We should note that Merton and Bodie (1995) in their work established that derivative instruments, traditionally not viewed as integral to the payment system, serve as an important extension of the payment system because they substitute in a variety of ways for trading in cash-market instruments.

Pooling of Resources

Firms use large amounts of capital provided by intermediaries that pool the savings of smaller investors. Traded assets, which do not provide funds for firms, but do establish the cost of those funds through market prices, also are purchased by pools such as mutual funds. Securitization of pools of mortgages, for example, transforms such pools into marketable securi-

ties. Again we see the standard short/long derivative, which underlies the nature of intermediation.

Transfer Economic Resources

Can a person buy future consumption? Can a society buy future consumption? A person does that by acquiring capital assets (even if these are statutory assets, such as Social Security, they are capital assets nevertheless). Society cannot just buy a lot of capital assets to assure future prosperity—especially if those are excessive claims to society's own cash flows. In the case of government bonds, this is known as *Ricardian Equivalence* (Brown 1995). David Ricardo, eminent classical British economist, believed that government bonds do not constitute real wealth because, in order for their value to be realized, government must extract resources from the economy in the future. Thus, Ricardo proclaimed that the government might as well impose taxes now, or, if it does not, other economic decision makers know that it might as well do so.

This idea has been recently given new life by the works of the ''Rational Expectations School'' in economics (Brown 1995). However, the society can prefund future consumption by facilitating economic investment. If we build factories now, cars will be produced in them in the future. What assures us that we will need those cars, or that the production will really happen? There is no complete assurance, but thanks to financial markets and financial intermediaries, we know what it costs to fund such economic projects, and we can evaluate if they are economically viable by comparing their net present value with the net present value of alternatives. Thus, an individual stores consumption by using capital assets, including those created by financial intermediaries, while the economy receives signals about cost of capital. It's the ''Amazing Invisible Hand'' at work—the greatest magician of all times.

Managing Risk

One does think immediately about insurance, but this also includes hedging and diversification (e.g., by mutual funds). This paper will study in detail the nature of the short/long position, which is at the core of financial intermediation, as it applies to the insurance industry. A life insurance company, for example, is long the stream of premiums, as specified in the policy, and short a very specific option on the human capital of the life insured—an option that becomes effective when the human capital value drops to zero, but whose exercise price is specified by the policy.

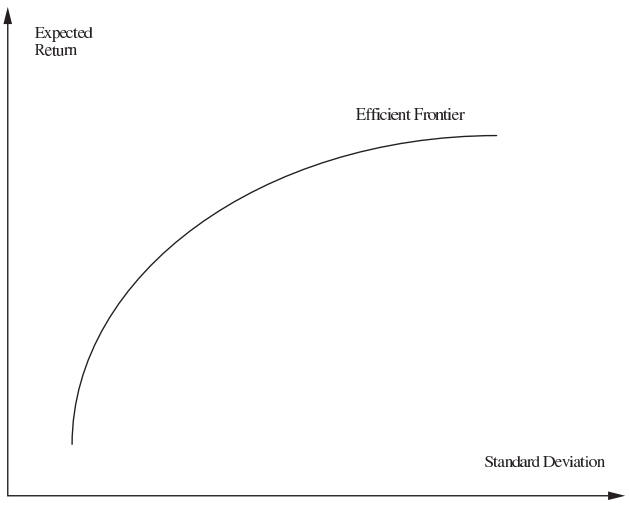
Similarly, a property-casualty company is long the premium stream and short certain contingent liabilities of the policyholder. Neither of these options is generally available in the market, because their risks are diversifiable, although the recent introduction of catastrophe futures provides a new perspective on the subject, which will be discussed later.

Providing Information

As noted earlier, financial markets, and financial intermediaries, by providing prices of capital assets, give signals to economic decision makers concerning costs of funding projects. It should be noted, however, that no capital assets are priced on the basis of the interest rate or yield curve alone. This was the founding principle of the Modern Portfolio Theory created by Markowitz (1952, see also Elton and Gruber 1987), when he noted that investors seek to maximize returns for a given level of risk. Assuming investors whose preferences are described by the mean-variance pair (i.e., expected return and the variance of the return), he then proceeded to note that investors will observe the efficient frontier consisting of all investments providing maximum expected return for a given variance level (see Figure 1).

Note that this efficient frontier is not only observed by the capital markets investors, but also by the firms

FIGURE 1
EFFICIENT FRONTIER



seeking financing of their projects. It acts as a ''price list of capital'' by identifying an expected return required by investors for a project at a given level of risk. Thus, financial intermediaries, by pricing derivatives created by them, help identify projects that rational firms should be funding.

As Merton and Bodie (1995) note, there is one more side to this story. Markets and intermediaries, by providing information about expected returns, implicitly inform economic decision makers of the risks of investments. The example given by Merton and Bodie relates to traded options markets, which have assumed an increasing role since their inception in 1973. The prices of options can convey implied volatility (i.e., standard deviation of the random return) of the underlying security, thus providing a risk measure derived from market signals.

Dealing with the Incentive Problems

One of the main examples of this is the problem of an optimal structuring of the relationship between the owners of an enterprise and the hired employers. This was referred to as *the agency problem* in the seminal paper on the issue by Ross (1973). A *principal* is a party that has controlling authority and that employs another party, called an *agent,* to act subject to the principal's control and instruction. The agency problem expresses itself in the uncertainty faced by the principal about whether the agent whose services are used will, in fact, act in the best interest of the principal and not in his or her own interest. Other such problems include moral hazard, adverse selection, and information asymmetries. A financial intermediary can create a derivative between the agent and the principal, and eliminate some of the friction. Again, Merton and Bodie (1995) discuss recent security innovations and the use of derivatives within corporate risk-management programs as examples of how the incentive problem can be resolved.

The previous discussion illustrates the main point of this paper—that the complicated maze of financial intermediation, and the philosophical perspective on the mission the insurance industry, can be simplified if we look at it from a distance and focus on *finding the derivative*. The derivative is generally created by the short/long position, with the creation on the short side of the trade. It is a private issue of the intermediary to its customer, not a marketable security (although a secondary market can develop for certificates of deposit, e.g., issued by banks). There also are marketable derivatives that serve a similar purpose to that of financial intermediaries. Merton and Bodie (1995) cite examples of the innovation spiral, which creates traded derivatives serving the same function as previously existing private issue derivatives (i.e., classical intermediaries' products). The implicit warning in their message is that a financial firm concentrating on the institutional perspective of markets, and not considering the functional perspective, may indeed miss the existence of its main competitors. Producers of carriages and buggy whips in the early 20th century most likely were not studying the efforts of Henry Ford.

It would seem appropriate to term the classical liabilities of financial intermediaries *private issue derivatives,* as opposed to traded derivatives. Interestingly, such a definition forces us to include over-the-counter swaps among the classical liabilities of financial intermediaries, and we believe this inclusion to be conceptually correct. The implication is that, if insurance firms are either long- or short-traded derivatives, they should value them the same way the market does. In contrast, private-issue derivatives must be given *market-related value,* as provided by modern financial mathematics, given the ''shadow competition'' of financial markets innovation.

Pricing Theories

This, in fact, brings us to the central issue in the integrated analysis of assets and liabilities. How should the prices of liabilities issued by financial intermediaries be determined? Asset portfolios of insurers generally are priced by capital markets. Pricing of liabilities traditionally has been the job of actuaries. But modern financial theory offers a perspective on this issue. There are two major theoretical approaches to pricing of claims to cash flows: the *no-arbitrage* and the *equilibrium* approach.

The No-Arbitrage Approach

The no-arbitrage approach can be used if there is a set of market prices for certain securities (i.e., marketable securities), and other securities can be replicated as portfolios of marketable securities. (Both contingent and deterministic claims to cash flows will be referred to here as *securities.*) This is, obviously, not always the case. Human capital, that is, a person's claim to his or her own earnings stream, is not mar-

ketable. It is additionally assumed that markets in which securities are traded are frictionless, with no trading costs, taxes, etc. Finally, the *principle of no arbitrage* holds, or, as economists would say, ''There is no free lunch.''

By the classical economic definition, ''arbitrage is the simultaneous purchase and sale of an item with an expectation of profit'' (Brown 1995). Profit exists if the same item has two different prices. The principle of no arbitrage requires that profits cannot be obtained this way by trading securities. Note that if one, indeed, could simultaneously sell an item for a higher price and purchase it for a lower price, a positive cash flow could be obtained with no payment for it, in other words, a free lunch. This principle therefore simply states that, if you plan to profit from a security, you must have paid for it.

Ross (1976) used this idea to develop a theory of no-arbitrage valuation of securities. The theory can be summarized as follows: If there are no arbitrage opportunities, then there is a positive linear operator, call it L, that can be used to value all marketed assets. Observe first that the set of all securities is a vector space, and the natural definition of a linear operator from linear algebra can be utilized. L is positive as it assigns nonnegative value to a stream of future nonnegative cash flow, with strictly positive values for nontrivial (not all zero) stream of cash flows. Finally, L assigns to a riskless pure discount bond its actual discounted market price. Once again the existence of the operator L is derived from the principle of no arbitrage. The operator is, in turn, given by a probability measure (called the risk-neutral measure or the equivalent martingale measure) with respect to which the price of a security is the expected discounted value (Dybvig and Ross 1987). This idea will be discussed further in Chapter 7.

As previously indicated, however, not all securities have their prices established by the market, and not all of them can be replicated by portfolios of marketable securities. In absence of these conditions, a more general economic approach must be utilized.

The Equilibrium Approach

The equilibrium approach looks at the agents in the economy. They are assumed to be rational wealth maximizers trading in the existing financial securities markets and subject to their resources' constraints. Equilibrium in the economy is obtained when no agent has any more incentive to trade, and the market clears. If the preferences of agents can be modeled, one can derive prices of claims to cash flows.

The comparison of the two approaches sheds light on the role of financial intermediaries. Banks and insurance firms issue securities whose cash flows are supported by cash flows of assets held by them. These newly created securities are nonmarketable, at least in the initial stages of their creation, and are derivatives. Yet, as markets for securities develop, these derivatives either become traded or can be duplicated by portfolios of marketable securities. This results in pricing by the equilibrium approach being replaced by the no-arbitrage approach. Could this mean that financial intermediaries eventually will be eliminated? Obviously, a necessary condition for that would be that all securities become marketable—a rather unlikely outcome, despite the broadening field of marketable claims.

There is more, however, to this issue. Not all market participants have the same position. An interesting perspective on this matter is given by Franke, Stapleton, and Subrahmanyam (1994). They explain that bundling of marketable derivatives, such as interest rate options (e.g., minimum interest rate guarantees) into insurance products may not be an outdated practice after all, even though such derivatives must be ''priced to market.'' They also ask a very fundamental question: ''Who buys and who sells options?''

Franke, Stapleton and Subrahmanyam's (1994) model is based on the equilibrium method. The authors assume a two-date economy with the dates indexed 0 (now) and 1 (future). There are I investors, $i = 1, 2, \ldots, I,$ in the economy. X is the risky payoff on the market portfolio at time 1. There is a complete market for marketable claims, so that each investor can buy state-contingent claims on X. Each investor chooses a payoff function $g_i(X)$, called the *sharing rule.* But, in addition to the investments in the marketable state-contingent claims, the investor also faces a noninsurable background risk, a random variable $e_i,$ so that the investor's total payoff at time i becomes:

$$y_i = g_i(X) + e_i.$$

The background risk cannot be traded. This results in the background risk affecting the optimal sharing rule applying to that investor. The investor solves the following optimization problem

$$\max_{gi(X)} E(u(g_i(X) + e_i))$$

subject to:

$$E((g_i(X) - g_i^0(X))\phi(X)) = 0,$$

where $u(w)$ is the investor's utility of wealth, $g_i^0(X)$ is the investor's endowment of claims on the market portfolio payoff X, and $\phi(X)$ is the *market pricing kernel* (in a complete discrete market, it is the price of a claim that pays a dollar in state X divided by the probability of occurrence of that state). In simple terms, the investor is aiming at maximizing expected payoff, subject to his ability to pay for the desired sharing rule at market prices. By assuming a hyperbolic absolute risk aversion (HARA) form of $u(w)$, i.e.,

$$u(w) = \frac{1 - \gamma}{\gamma}\left(\frac{A + w}{1 - \gamma}\right)^{\gamma}, \qquad (1.1)$$

and the same form of this function for all investors, but with different levels of the variance of e_i, Franke, Stapleton and Subrahmanyam (1994) show that the optimal sharing rule of agent i is

$$g_i(X) = A_i^* + \alpha_i X + \alpha_i(\psi_i^*(X) - \psi(X)), \quad (1.2)$$

where A_i^* is that investor's risk-free income at time 1, $\alpha_i X$ is that investor's linear share of the market portfolio payoff, and $\alpha_i(\psi^*_i(X) - \psi(X))$ is that investor's payoff from contingent claims. If there were no background risk for all agents in the economy, all investors would have only the first two terms in their sharing rule, which are linear, while the third term is caused by the background risk and is nonlinear. Thus, the presence of background risk produces a nonlinear demand for claims on the market portfolio. An agent with relatively high variance of background risk will have a relatively convex third term, and this will result in that agent purchasing contingent claims. An agent with no, or a very low level of, background risk will have a concave third term of the sharing rule, and will sell contingent claims in the market.

The conclusions of Franke, Stapleton, and Subrahmanyam (1994) are of significance to the insurance industry. By the reasoning presented earlier in this chapter, financial intermediaries provide a variety of contingent claims backed by holding a portfolio much more closely resembling one available in the market. The model suggests that this is simply because insurance firms face significantly less background risk than their customers do (with similar conclusions obviously holding for other financial intermediaries).

It should also be noted that an essential part of this model is the distinction between marketable securities and the nonmarketable background. Such a distinction is clear at any given point in time. However, over time, more and more financial claims become traded assets. A recent introduction of futures contracts for delivery of electricity or catastrophes are examples of such a process.