



International
Financial
Risk Institute

Geneva,
Switzerland

Bank Capital and Risk Management: Issues for Banks and Regulators

By Prof. Kenneth A. Froot, Harvard University and NBER

7th Risk Management Round Table 5 & 6 April 2001

IFCI Geneva Research Paper N° 8

*IFCI Geneva Papers are research publications sponsored by the International Financial Risk Institute (IFCI), 2 Cours de Rive, CH-1204 Genève,
Tel. ++41-22-312.56.78, Fax. ++41-22-312.56.77.*

*Information on the Institute as well as on the IFCI Geneva Papers are also available on the World Wide Web: <http://www.riskinstitute.ch>, e-mail: info@ifci.ch.
©IFCI (reprinting or quotations citing source permitted).*

Bank Capital and Risk Management: Issues for Banks and Regulators¹

Kenneth A. Froot
Harvard University and NBER

June 2001

Abstract

Banks and financial firms are in the process of evolving away from primary warehouseers of risk to diversified originators and distributors of financial services. These changes are important for the way that financial firms think about their needs for economic capital and capacity to bear risk. They are also important for the way regulators evaluate capital and connect it to concepts of systemic and liquidity risk.

This essay tries to grapple with the implications of such changes. I argue that the BIS standards place important, but perhaps excessive, emphasis on risks that emerge from financial contracts that warehouse risk and pay insufficient attention to non-warehouse business risks that emerge from broader bank activities. I also argue that the profitability of non-warehouse businesses alters economic capital requirements and necessitates changes in standard VaR calculations, and that non-warehouse businesses create important externalities in regulating systemic and liquidity risks. I believe that the Basel Committee's proposals, while providing a significant move forward, lag behind these developments in financial firm activities and may lead to considerable regulatory distortions as a result.

¹ Prepared for the IFCI Risk Round Table, Geneva, Switzerland, April 5 & 6, 2001. Many thanks to Sid Browne, Richard Evans, Stephen Kealhofer, Mark Kritzman, Bob Litzenberger, Charles Monet, Tim Shephard-Walwyn, Thomas Wilson, and Timothy Wilson for stimulating conversations and suggested readings. All errors remain my own.

capital that relies substantially on book value and on definitions of risk that emanate from balance sheets, rather than from total business risk.

I. Definition of Risk and Capital in Modern Financial Institutions

As financial institutions evolve, the definition of capital needs to evolve as well. To have a meaningful discussion of capital requirements for today's forward looking financial firms, we need to clarify and develop the meaning of capital in financial firms.

Historically, the primary purpose of financial firms – be they commercial banks, investment banks, insurers or others – was to use their balance sheets to intermediate in relatively illiquid assets. Illiquid assets are informationally intensive. They may be subject to potentially important informational asymmetries in their trading, or they may lack standardization, and therefore require a high degree of individual inspection. Classic examples of such illiquid assets include small and mid-size company loans, all forms of insurance contracts, young-company equity and debt claims, real estate loans, etc.

Traditionally, it is through the intermediation process that illiquid assets have become more liquid. Banks and other financial institutions aggregate exposures, each of which is extremely illiquid, and finance the exposure pool through the issuance of their own liabilities and equity. The liquidity of these pooled claims against banks is far greater than the liquidity of the average exposure that the bank takes on. In this way, financial firms reduce transaction costs in comparison with a system that passes each exposure directly into the financial marketplace.

However, while traditional financial firms reduce transaction costs associated with direct market access, by no means do they minimize the transaction costs of warehousing. Indeed, when compared with special purpose vehicles, financial firms may actually *subtract* liquidity in the warehousing of assets they originate. There are a number of reasons for this.

First, pools of claims are not very transparent when held by banks. Often, little in the way of specifics is provided on the composition of the portfolio of assets or on the risk characteristics of that portfolio. This is not the case in many special purpose vehicles that provide investors with considerable detail about the composition, historic performance, and risk characteristics of pooled assets.

Second, inside financial firms, the pools change over time according to the discretion of managers. Managers receive salaries, perks, bonuses, equity, and options and are therefore not likely to have incentives that are perfectly aligned with those of any specific capital provider and certainly not those of equity holders. In addition, for a variety of reasons, managers do not provide details on how their firms' asset pools will evolve over time. Such dynamic discretion is especially important for financial firms, since their risk balance sheet profile can change far more quickly than it can for nonfinancial firms. Many special purpose vehicles have clearly-defined rules circumscribing dynamic managerial discretion, sometimes eliminating it completely.

Third, capital providers historically have no choice but to fund the pool of assets (i.e., the warehouse) together with the associated origination and distribution activities. Banks make loans and take on credit exposures, insurance companies warehouse their underwriting opportunities, etc. In some instances it may be more profitable to separate the warehousing from the origination and distribution. As I discuss below, this contributes to considerable inefficiency and lack of information production about the asset pool and about financial firm performance.

Fourth, traditional financial corporations have a tax status that resembles other corporations, rather than that of pass-through vehicles that are not subject to corporate taxes. This adds an extra, but unnecessary, layer of taxation to the warehousing function when it is located within a standard bank or insurance company.

Finally, traditional financial firms are highly regulated. In some instances regulation may be a justifiable cost – it makes financial firms behave more prudently than they otherwise would. However, in other circumstances regulations may make firms behave more conservatively than markets realistically require. By asking markets to provide collateral dedicated to specific warehousing portfolios it is possible to use market prices to gain a better gauge on the appropriate level of conservatism at the same time as shifting any implied performance guarantees away from governmental institutions and towards investors.

All these factors contribute to additional costs of warehousing in traditional financial firms, costs for which capital providers necessarily charge. A higher cost of capital makes the warehousing function of financial institutions relatively inefficient.

In recent years, there has been an enormous increase in the rate at which forward-thinking financial institutions have been responding to these inefficiencies. Essentially, they have been diversifying into fee businesses associated with origination and distribution, at the same time as they have trimming their traditional warehousing functions. Securitization of assets and asset pools and the creation of special purpose vehicles to warehouse them occurred on a wide scale, and will continue for many financial institutions over the next decade. In addition, there has been considerable growth in off-shore activities which are designed to provide warehousing in tax-advantaged and friendly regulatory environments.

Securitized pools in principal avoid many of the inefficiencies mentioned above. They exhibit far greater transparency – both static and dynamic – for their capital providers. Given their dedicated, special purpose nature, they provide for little or no managerial discretion, and therefore far less opacity. Furthermore, there is typically no layer of corporate taxation, since such entities are mere pass-through vehicles.

This change in the mix of financial firm activities has far reaching implications for both the capital markets and the definitions of economic and regulatory capital at financial firms.

First, for capital markets, the dynamics change the reliance on markets versus institutions to deliver liquidity. For example, for claims that were warehoused by banks, liquidity arose through the central bank providing funds to financial firms to avoid their needing to call in loans. Liquidity was in part provided through the use of deposit insurance which, by guaranteeing depositor liquidity, stemmed the bank run phenomenon. However, even with deposit insurance, banks under financial pressure reduce liquidity by failing to lend, as evidenced by Japan over the last decade.

Second, the proliferation of arms' length capital market transactions between originators and warehousemen changes the norm surrounding book versus market valuation. If portfolios of loans that back CLOs are implicitly valued in the market on a daily basis through the daily pricing of the CLOs, why shouldn't one do the same for portfolios of loan that back bank liabilities? And if one marks to market the assets of banks, wouldn't we use market values for measuring bank capital?

A. Defining Capital in Financial Firms

This latter effect is a completely positive one. Capital in financial institutions can only be sensibly defined to include the difference between the market value of the assets and the default-free value of customer liabilities. The distinction between customer and investor liabilities is crucial. Any loss to customer liabilities can be a serious impediment to confidence in a financial firm. Customers do not wish to bear performance risk or to be paid as an investor would for doing so. Thus, it is imperative for the ongoing performance of a financial firm to assume that these customer liabilities are honored.

The treatment of investor liabilities is more complicated. Some argue (see, for example, Shephard-Walwyn and Rohner²) that for the purpose of establishing minimum capital requirements, only equity or equity-like claims that are able to sustain losses without any encumbrance ought to be counted. Thus, under such an arrangement, capital would be reduced by the default-free value of investor liabilities as well as by the default-free value of customer liabilities.

Such a definition of capital is appropriate to delineate a buffer to protect against stress losses, particularly if they are systemic in nature. Risk transfer without default is a critical aspect of the funding to withstand such shocks. In the strongest case, this would argue in favor of using the default-free value – not the market value – of investor liabilities. In stress environments when bank liquidity and funding is threatened, the likelihood of default on investor liabilities contributes to the firm's illiquidity and stress, and therefore should reduce capital. If one were to use the market value, rather than the default-free value of investor liabilities, a firm could show considerable positive capital, yet have its capital funding in jeopardy, even if its customer liabilities were reasonably sure to be paid.

² Tim Shephard-Walwyn and Marcel Rohner, "Equity at Risk: An Alternative Approach to Setting Regulatory Capital Standards for Internationally Active Financial Firms," IFCI, April 2000.

The distinction between using the default-free value versus market value of investor liabilities may seem like a small one. Default is rare enough, and in any case investor liabilities would not constitute a majority of the difference between market-value assets and book-value customer liabilities. However, this is not the case. Indeed, by subtracting the market value of investor liabilities one is essentially counting investor liabilities as capital. That is because, in increasingly severe stress scenarios, the market value of investor liabilities goes to zero before customer liabilities are affected.

This is not to say, however, that subordinated debt can't – or shouldn't – be used as capital for some purposes. Indeed, a different argument can be made for other prudential regulations that seek to ensure a bank is sufficiently conservatively funded to protect its customers and deposit insurers. It is potentially useful to encourage, and even require, banks to rely on subordinated debt financing. As the US Shadow Financial Regulatory Committee (SFRC) has pointed out, the required use of subordinated debt would provide valuable information to the market and to bank supervisors.³ With subordinated debt present, analysts and market participants would pay additional attention to the risks of this balance sheet tranche. The price of subordinated debt in the market place would serve as a useful objective indicator of risk individual bank risk.

This type of information is important today, and will be increasingly important in the future. Regulators and bank supervisors could benefit, as sub-debt prices would provide a tangible performance benchmark and early warning signal. Furthermore, we are rapidly moving toward a world in which deposit insurance is priced according to the risk characteristics of pools of firms, and, eventually, to individual firms. For example, the US FDIC is currently in the process of developing a plan under which it would purchase deposit insurance in the marketplace (purchased from some combination of capital markets and reinsurers). In this world, sub-debt prices would reinforce pricing of deposit insurance and vice versa, leading to better overall information accumulation and greater incentives for bank transparency.

In addition to providing information, subordinated debt could play a disciplinary role as well. The interests of subordinated debt holders and deposit insurers are closely related, and the stakes of these claimants receive more attention with management once there is a capital-markets as well as a regulatory representation.

While subordinated debt provides substantial externalities, it is important to keep these in perspective. The SFRC takes the view that some minimum amount should be required and that there should be no upper limit as to the fraction of capital provided by subordinated debt, much as with non-financial firms. Clearly this generates concerns that individual bank failures and more generalized systemic incidents could be triggered as a result of the presence of the sub-debt. To ameliorate this possibility, the SFRC suggests that the issuing bank, at the direction of its supervisor, might withhold payment of

³ "Reforming Bank Capital Regulation: A Proposal by the U.S. Shadow Financial Regulatory Committee," George G. Kaufman, Robert E. Litan, Richard C. Aspinwall, George J. Benston, Charles W. Calomiris, Franklin R. Edwards, Scott E. Harrington, Richard J. Herring, Paul M. Horvitz, Roberta Romano, Hal S. Scott, Kenneth E. Scott, Peter J. Wallison. AEI Press, Washington, D.C., March 2000.

interest and principal if capital fell below a designated threshold. This is reminiscent of the way that 'surplus notes' work in the US insurance regulatory context.

This kind of discretionary 'out' substantially undermines the above information externalities that sub-debt-as-capital could otherwise bring. Rather than provide information about creditworthiness, price changes might be attributed to perceived fluctuations in supervisors' willingness to provide relief. Moreover, once the hard penalty of default is removed, there is less incentive for management either to provide information or to tailor bank policies toward protecting sub-debt claims. As in the case of US insurance, the primary benefit of this type of arrangement would be to provide the industry with a source of tax-deductible debt service. This benefit may be substantial. But flexible sub-debt could not provide information about creditworthiness to regulators or capital markets. Nor could it align interests of debt holders with depositors. As a result, I believe sub-debt is useful only to the extent that it can remain as true, inflexible debt, and it would therefore need to play a restricted role in constituting regulatory capital.

B. Profitability, Risk, and Capital

Given the evolution of financial institutions away from basic warehousing functions, the notion of 'capital' becomes more involved. Functionally, it is useful to think of capital as a buffer against negative shocks to firm value. In a world in which warehousing is overwhelmingly the most important function, a complete definition of capital would be as above – the market value of the tangible assets less the liabilities. In such a world, the regulator is perhaps appropriately focused on measuring the risks associated with warehoused exposures – market, credit, liquidity risk, etc., all take center stage.

However, in a world where warehoused assets may turnover – and rapidly – and where rents are earned from origination and distribution, tangible assets do not constitute the entire market value of the firm. There may be profits that accrue to dynamic portfolio management and origination and related service activities. If such profits are important, why wouldn't we consider them as part of the buffer against which loss events might be tallied? After all, if a loss is offset by profits, then balance sheet capital need not be called upon.

Reciprocally, if trading, origination and related service activities become important as sources of income, they may also become important as sources of firm-wide risk. Firms may experience stress events because of losses in these activities. So isn't it important to include these risks in measures of economic and regulatory capital?

With this in mind, recent papers have therefore turned toward expanding the concept of the buffer beyond traditional measures of balance sheet capital to include earnings and profits.

1. Earnings risk

There are a few important points to make about these views. The first is that it is entirely prudent to include non-balance sheet risks in concepts of economic and regulatory capital. It is appropriate, however, only once the decision has been made to combine the warehousing and origination/distribution functions in the same firm.

For economic capital, the case seems fairly unambiguous. Financial firms should consider their total business risks when thinking about economic capital. They should not confine themselves to merely the risks of their warehoused exposures, such as credit, market, liquidity, and operational risks.

There is abundant evidence that non-financial firms set their capital structures with an eye to such risks, though they often employ no formal risk measurement procedure. Business risks drive decisions about the retention of realized profits, as well as the composition of capital (debt versus equity) on the balance sheet.

A good example of this process at work in a non-financial firm comes from Microsoft. As is well known, Microsoft has locked in the operating system in approximately 90% of the installed base of hundreds of millions of PCs. It has (as of the end of its last fiscal year, June 2000) almost \$10 billion in annual free cash flow, a market capitalization of over \$400 billion, no short- or long-term debt, and cash holdings in excess of \$23 billion plus an additional \$17 billion in financial investments which together represent approximately 80% of its total assets.

At first glance, it seems like the firm holds an extraordinarily conservative amount of economic capital. The incremental annual deadweight cost from the taxes alone of holding this cash position is approximately \$450 million. The firm is just releasing its Windows 2000 operating system, and its products are shipped with most new PCs. The cash plus investments position has doubled over the last three years. Virtually any formally modeled analysis of the cash flow stream would suggest large positive inflows for years to come. To the extent that any firm in the software industry looks like a Treasury bill, it is probably Microsoft.

Indeed, a number of observers have suggested that Microsoft could easily borrow \$25 billion without a question of whether it was running out of debt capacity, saving another \$450 million or so in annual taxes. In this sense the firm has a negative leverage position of almost \$50 billion. This is a strikingly large amount of spare debt capacity.

However, the view among Microsoft's top management is that this position is not at all conservative. They view the \$900 million total annual tax cost from the negative leverage as a low insurance premium. It is Bill Gate's personal view that, given the business risks the firm faces, Microsoft needs to have a capitalization that allows it to run with *zero revenues* for one year. Indeed, in June 2000, Microsoft's revenues were \$23 billion and gross profits (revenues less costs of goods sold) totaled \$21 billion. Needless, to say, even with the recent US antitrust case, zero revenues for a year is a very low probability event. At current levels, Microsoft could run for *two years* without any revenue.

Obviously, the Microsoft example does not suggest that financial firms should hold comparable levels of capital against their business risks. For a variety of reasons, there may be far less pressure exerted by Microsoft's shareholders to conserve on internal capital, especially in view of Microsoft's historic performance and continuity of management. This lack of pressure may be good (Gates is expected to lead the firm to earn and then disburse even more funds later) or bad (the insurance is unnecessary at best and a license to make big mistakes at worst). But given the high returns on capital historically earned by Microsoft, running out of money would have been a very costly event.

Nevertheless, financial firms should be among the more conservatively structured in this regard. Business risks, in addition to the risks of the balance sheet, need to be modeled and included in the discussion of economic capital.

2. Earnings drift

Business risk concerns deviations of cash flow from expected levels. If we include unexpected cashflow in computations of capital, shouldn't we also include expected cashflow? It seems sensible that a highly profitable firm would need less capital than a twin firm with equivalent risk but lower profits. After all, the probability of a given size end-of-quarter loss will always be lower for the more profitable firm.

Unfortunately, it is not this simple. Risk grows at the square root of time while expected profits accrete linearly with time. The ratio of risk to expected return generally converges to infinity for time short intervals. As a result, the importance of earnings drift varies with the choice of horizon.

If all that mattered were the end-of-quarter results, then it would be this simple. But horizons in most risk computations are the least interesting part – horizons are really more of a standardization tool than anything else. Regardless of the horizon of a particular bottom-up VaR computation, what we really care about measuring is instantaneous risk. There are two reasons for this. First, we do not honestly think that the portfolio will remain constant and fixed over any reported horizon longer than a week or so (e.g., two-weeks, one month, etc.). Over time the actual distribution of the P&L will differ from the VaR because of intervening portfolio adjustments.

Second, and more important, is that the probability of breaching a given loss amount over a fixed period can be far higher than the probability of breaching that loss amount at the end of the period. Equity prices during 1987, the year of the crash, were unremarkable if one looks only at annual US returns. One could not guess how frightening and risky that Monday in October was given only the way things turned out. This is the first passage time problem.

First passage times are only rarely discussed in conventional VaR calculations based purely on riskless drift. It's not that first passage times aren't important for such

calculations – they can be material. But most people know how to calibrate the correspondence between a standard VaR and a VaR that measures the probability of crossing the threshold at any time over the horizon. Once we introduce drift – non-zero expected profits – the relationship between these two VaR calculations varies dramatically based on expected profits.

To see all this, it's best to consider an example. Let us take a standard VaR computation for a portfolio with a certain risk and expected return over a given horizon. The VaR merely tells us the losses on the portfolio that we expect to exceed a given percentage of the time at the end of the horizon. For instance, suppose a portfolio has a continuous expected return of zero and annualized standard deviation of 18% (roughly the historical average risk of the US stock market). A 95% VaR over a 1 day horizon would yield 1.54%, meaning that 95% of the time, we would expect the portfolio's return to exceed a 1.54% loss over the next day.⁴

Next, let us calculate a VaR based on first passage time. This VaR will give us the loss on the portfolio that we expect to exceed a given percentage of the time from inception to any point during the specified horizon. One can think of this calculation as looking across portfolio paths and finding the value of the portfolio that we expected to exceed continuously. To continue with the same portfolio example, a 95% VaR based on first passage time over a 1-day horizon yields 1.83%. In other words, 95% of the time during that the day, the portfolio value was above a loss of 1.83%. This result shows greater dispersion than the end-of-day VaR because extreme end-of-day losses are quite likely to have been preceded by even more extreme losses during the day.

⁴ Because this is only an example, I assumed volatility is evenly spread across a 365-day year. As is well known, the volatility during days that the market is open is much higher than the days on which it is closed, so that this result understates the risk on a day when the market is open.

Percentage By Which First Passage Time VaR Exceeds Standard VaR

	Horizon		
	1 day	1 quarter	1 year
Expected return = 0	19%	18%	16%
Expected return = 20%	20%	37%	104%

This table shows the sensitivity of VaR calculations to horizon. If we simply focus on risk, and leave expected returns out, the first line shows that there is a 15%-20% deviation between standard VaR measures and first passage time VaR. For example, the ratio for 1 day with a continuously expected return of zero is just $19\% = 1.83\% / 1.54\% - 1$. Increases in risk on the order of 19% are not insignificant. However, as the first line of the table shows, the increase is not very sensitive to horizon, so it is not particularly troublesome.

However, when expected returns are non-zero, the difference between these two risk measures becomes very sensitive to horizon. As we already know, over short periods of time, expected returns don't matter for the risk computation – over a 1 day period, a 20% annual expected return is less than 10 basis points. So there is little effect on the VaR. However, at longer horizons, the two risk measures begin to diverge considerably. Indeed, over very long (e.g., 3 year) periods of time, the ratio computed in the table no longer makes sense. With a 20% expected return, the standard VaR shows a *positive* return, since the 95% level of the distribution is an absolute gain. However, this is not true for the first passage time portfolio, which continues to show large losses at 95%.

How can we make sense of these numbers? Of course, we don't literally care about first passage time VaR. Such a computation is most useful if there is an inflexible and meaningful threshold that we don't want to cross, trigger perhaps by a debt covenant, or a particular level of losses. No firm is literally going to be shut down because of a momentary breach of economic or regulatory capital requirements.

However, first passage time is in the spirit of what we are trying to achieve with capital measures. This is particularly the case because much of the concern with capital standards is driven by customer and rating agency expectations rather than solvency per se. One cannot know the point at which questions begin to be raised by customers and/or regulators about bank viability and liquidity. But surely that point matters for determining capital, and just as surely, it is not driven by end-of-period values.

With this in mind, we can return to our motivating question in this section: shouldn't we have lighter capital requirements for a firm with higher profits but the same risks? If we take the first passage time concept seriously, we still see an impact at longer horizons of profitability on risk. The table below shows how first passage time VaR changes with expected return at different horizons. Naturally, there is no impact of expected returns at

short horizons. But at quarterly intervals, there remains a substantial reduction in risk as expected profitability increases. This reduction is much smaller than under standard VaR measures, but it is noticeable.

It is interesting to note that at 1 quarter, the standard VaR computation used in Table 1 above yields 13.7% and 9.3% for expected returns of zero and 20%, respectively. Thus the standard VaR result with zero expected returns is essentially the same as the first passage time VaR result with a 20% return, which the table below reveals to be 12.8%. In this example, therefore, adding expected profits reduces losses by about the same amount as account for first passage time increases losses.

First Passage Time VaR			
	Horizon		
	1 day	1 quarter	1 year
Expected return = 0	1.8%	16.2%	29.7%
Expected return = 20%	1.8%	12.8%	18.6%

Before leaving the computation of risk issue, recall that by necessity we continue to assume that the portfolio of risks remains constant throughout the entire horizon and that the portfolio returns are instantaneously normal with constant mean and variance. As mentioned above, these approximations are likely to become greater sources of concern at horizons like a quarter or a year.

The important message from this section can be summarized as follows. Standard VaR measures basically estimate continuous risk over a short interval. Risk could be measured this way over longer intervals (e.g., 1 quarter or longer), but that would inappropriately emphasize the importance of an arbitrary an end-of-horizon date. As a result, expected returns – which matter only at longer horizons – don't have a very important practical impact on VaR calculations. First passage times, however, are concerned with the entire horizon and therefore do not emphasize the end-of-horizon date so much. So first passage time probabilities are more appropriate for gauging risk over longer horizons when expected returns are material. Efficient capital markets may drive expected excess returns under risk neutral probabilities to zero. But efficient product markets in non-warehousing functions, such as origination, distribution, and servicing, are much less likely. Our measures of first passage times show that non-warehousing businesses can have a major impact on economic capital requirements of financial firms.

C. Asset Values and Firm Collateral on Call

The above discussion is helpful for clarifying the measurement of risks in financial businesses with non-warehousing functions. What should the relationship be between these risks and the market value of the firm's capital? Non-warehouse businesses often require little book capital, but can have substantial market value. Can we associate this market value with the economic capital required to support the risks? After we have

measured risk, we need to pair that risk against some measure of economic capital. Once again, adding firm operating profits to the equation requires some work to create a consistent definition of capital.

The basic point is twofold. First, the sources of value within the firm can be expressed either as the value of tangible assets or as the value of intangible assets. We need to make this distinction to know what it is we have been counting as capital all along. Second, these sources of value are only useful for capital to the extent that they can be turned into cash. We will therefore have to struggle with the collateralizability of the sources of financial firm value. This is what is meant by firm 'collateral on call.'

1. The value of balance sheet assets

To start off, we must ensure there is no double counting and that we have a clean separation of value. That is, we need to be clear about which profits are counted in the market value of assets and which are not.

To see why this obvious point matters, suppose a bank has made an infinite maturity loan (a perpetuity) of 100 at a 10% spread without incurring any credit risk. If discount rates are 5%, the market value of this asset $100 + 10/0.05 = 300$ and the net contribution to capital from this asset is $10/0.05 = 200$. The important point is that the spread – the future stream of profits associated with this loan – is already capitalized in our market-value measure of capital. Anyone looking at the assets of the bank would place the market value of the loan at 300. It would therefore be a mistake to count the future profits as a buffer *over and above* the market value of the assets.

However, if we change the story slightly, we get a completely different answer. Suppose that the bank has the same riskless opportunity to lend 100 each year at a 10% spread, again with no credit risk. With discount rates at 5%, the contribution to capital of the going concern activity is again 200. However, if the loan agreement is structured as a one-year note to be rolled, rather than as a perpetuity, a regulator or a risk manager looking at the exposure warehouse would reasonably decide that the market value of the asset is just 110. As a result, the contribution to capital is only 10 – the value of the profits on this year's loan only. In this case, it would be necessary to count the remaining 190 of future profits as a buffer that is *additional* to the market value of the tangible assets.

Thus, our first point about including profits in an expanded definition of capital is that dynamic changes in the existing portfolio have a value (positive or negative), and that we will miss these if we simply add up the value of the identifiable assets in place. If we were to use the market value of existing assets as a complete measure of the going-concern value of the exposure warehouse, we would need to assume that future changes in the portfolio will always be executed at NPV-zero values. Clearly, in some instances this assumption will be wrong – the market value of the firm is higher than the market value of the assets in place.

The corollary of this is that origination, distribution and all other business activities whose value also cannot be captured by the tangible warehouse assets also will not show up in measures of warehouse capital, even when market values of the warehouse assets are used.

This logic obviously pushes in the direction of using the market value of the firm as the basis for measuring capital, rather than the value of some set of tangible assets. So if we are to advocate including future earnings in an expanded definition of capital, does that mean that we wind up defining the entire present value of the firm less its liabilities as capital?

2. Collateral on call

This brings us to the second issue in this section: the collateralizability of sources of financial firm value.

In a perfect market with complete information and contracts, there would be no distinction between value and collateral. The value of virtually any asset could be realized quickly in the marketplace. In this world, our definition of capital would be the market value of the firm less its liabilities.

It is important that perfect collateralizability is a more comprehensive notion than perfect liquidity, and therefore much harder to achieve. For example, suppose a publicly-traded firm sells additional shares. Prices may become temporarily depressed because it takes time to assemble buyers with incremental demand at incipiently lower prices; this is what we mean when we say liquidity is not perfect. However, something else happens when an established firm sells shares: shareholders receive a negative signal about true value. Because managers tend to sell when their private information suggests overvaluation, an announcement of an equity sale is negative information the market previously didn't have. The result is that share prices fall. The pre-issuance share price is not achievable, and therefore the firm's market value is not collateralizable. The larger the offering, the larger the price decline. This is true even if the shares are perfectly liquid in the usual sense.

Note that equity-issue price declines do not have to be very large to substantially reduce collateral on call. For example, suppose that an ongoing firm has investment opportunities but currently is low on internal funds. It has a market value of 1000 and needs to raise 100 to fund the opportunities. If the price decline is 5% (which is typical in information intensive industries for this size equity raise), then the market value of the firm including the additional 100 is 1050 ($1000 - 50 + 100$). The firm issued 100 in new shares, but increased its total value by only 50. Thus, the cost of collateralizing market equity is 50% for this firm.

The degree of collateralizability naturally differs across sources of value. The assets in place (e.g., loans) may be illiquid, but they can often be collateralized at close to fair value. This is especially true if the central bank is prepared to purchase such assets

directly. But even in the absence of central bank intervention, there are liquid markets for loan pools. Evidence for this comes from CLOs, where pools of loans are sold to special-purpose financing vehicles at competitive prices. Other, less tangible assets cannot be so readily collateralized. The seller of claims to highly intangible assets, such as a firm's growth opportunities, faces the same information and control problems as the firm that announces an equity offering. Growth opportunities cannot be sold easily, because their capture involves an intricate combination of inputs, including specific individuals, knowledge, information, relationships, etc.

Much academic work has sought to put more structure around this thinking.⁵ Because there is asymmetric control over and information about future profits, it is expensive to collateralize them. Firms will find it costly to raise external funds in large amounts based on such opportunities. The costs are greater if the opportunities are opaque and if managerial discretion is greater and smaller if they are transparent and contractually perfected.

To continue with the example of the bank that lends at a riskless 10% spread, the market might well value rolling the one-year note spread at full net price – 200 if the opportunity were funded internally. However, if the firm were to try to raise external equity to fund the loan, the equity would sell likely at a lower price. Information asymmetries and adverse selection make shareholders skeptical that managers see the prospects as good. Of course, the closer that the bank can come to demonstrating publicly that the loan must be rolled, the more the profits become more transparent and mechanical and the more cheaply they can be collateralized.

This logic forms the basis of theories of risk management and capital allocation, applied both to financial and nonfinancial firms.⁶ In these theories, firms have “internal” funds (e.g., reasonably liquid assets in place) that can be used to make investments or cover losses. Firms can gain access to external funds, but there is a cost to doing so, since that requires them to collateralize informationally-intensive profit opportunities. Additional risks that they take on must be priced according to how much additional costly collateralization will be needed.

Clearly, we cannot include the market value of all firm assets in a sensible definition of economic or regulatory capital. Transformation of the market value of the firm into internal funds involves a reduction in value. This value reduction is low if a well-capitalized firm borrows short-term funds, and it is large if a firm with little financial slack needs to raise a considerable amount of equity. Many internet companies in 2001

⁵ For a survey of the issues and literature, see Froot, K., “Incentive Problems in Financial Contracting: Impacts on Corporate Financing, Investment, and Risk Management Policies,” in *The Global Financial System: A Functional Perspective*, Harvard Business School Press, Boston, 1995.

⁶ See Froot, K. and J. Stein, “A New Approach to Capital Budgeting for Financial Institutions,” *Journal of Applied Corporate Finance*, Summer 1998b, 59-69; and “Risk Management, Capital Budgeting and Capital Structure Policy for Financial Institutions: An Integrated Approach,” *Journal of Financial Economics*, 47, January 1998a, 55-82. Also see Froot, K., D. Scharfstein and J. Stein “Risk Management: Coordinating Corporate Investment and Financing Decisions,” *Journal of Finance*, 48, December 1993, 1629-1658 and “A Framework for Risk Management,” *Harvard Business Review*, 72, September-October 1994, 59-71.

are an extreme example of the latter, with positive market value but access to external funds only at near-infinite costs.

What about using only next-quarter profits as opposed to the entire future stream of discounted profits? Once again, we have to ask how easily such profits can be collateralized. Clearly, if the firm has untapped debt capacity then such profits might be financed without much cost. However, if the firm is run thoughtfully, it should not have untapped debt capacity. Even in the extreme example of Microsoft above, we should probably view the excess debt capacity as dedicated insurance collateral. If that case, even Microsoft would endure important costs in collateralizing one-quarter ahead profits. Drawdown of insurance collateral depletes the firm's financial slack.

To conclude this section: firm market value in excess of warehouse-asset market value should not be immediately counted as capital for economic and regulatory purposes. Indeed, a conservative view would not count any of the excess as capital. In fact, a conservative view would also take into account the liquidity risk of the warehouse assets, thereby further reducing capital. In individual instances, of course, such a conservative view may not be justified: a portion of the firm-warehouse market excess might be collateralizable. Unfortunately, however, there is no easy way to determine generally the degree to which collateralization is feasible.^{7, 8}

II. The Internal Pricing of Risks

Once we have thought about measuring both risk and capital, the next most important issue is concerns pricing: what is the cost of the capital being deployed? When can we feasibly use a pricing system to help allocate capital?

Our discussion of these issues will proceed in two parts. First, we need to be clear about the building blocks of capital pricing and allocation for the more traditional warehouse case in which risks are driven by balance sheet exposures. This is an area where there has been a tremendous foundation built, both conceptual and practical, at major firms. There are a variety of approaches across banks, and I won't attempt to summarize them here. Rather my goal is to layout a more conceptual framework based on a few building blocks.

The second part of the discussion is to ask how we should modify things for financial firms that are more focused on origination and distribution. But at first, we focus only on the warehousing functions.

A. Pricing Capital in Warehousing Activities

⁷ We still also face the issue of illiquidity of balance-sheet assets. See Stephen Kealhofer, "Liquidity, Liquidity Crises and Bank Capital Regulation," discussion paper, 2001 for a discussion of these issues.

⁸ In the above VaR simulations, we counted profits as they were received, and we assumed that in expectation, profits accrete continuously over time.

The main conceptual building blocks in a system of capital allocation and pricing are as follows:

1. Capital allocation *is not* capital pricing
2. Prices pay capital providers for systematic risk exposures
3. Prices pay capital providers for corporate financing imperfections
4. Prices pay capital providers for intra-corporate financing imperfections

Together these lead to a system that both prices capital in various uses and determines its allocation. We take them one at a time.

1. Capital allocation *is not* capital pricing

Many internal systems rely on the logic of distributing physical units of capital rather than risk. This made sense until about 15 or 20 years ago, when in banking buying a risk generally necessitated relinquishing cash – i.e., making a loan, buying a stock or note, etc. Since that time, the growth of derivatives has made it a commonplace to separate financing and risk. For capital markets risks, it has swiftly become the case that it is risk, not capital that needs to be allocated. Through good pricing, risks are taken to maximize value and capital is there to provide a buffer.

Even though this separation between financing and risk has caught on quickly, is not perfect. Exchange-traded futures, for example, require variation margin so that the exchange does not have to bear the credit risk of every contract holder. Conventions like variation margin, overcollateralization, repo haircuts and the like are evidence that financing and risk are not fully independent. But in many circumstances the dependence can be avoided. Large firms, for example, can easily avoid variation margin by careful structuring (e.g., using a forward instead of a future) and through credit risk transactions. So, for the purposes at hand, the separation between financing and risk taking is a legitimate starting point for capital associated with bank warehouse functions.

Under these circumstances, capital is a risk buffer and shouldn't be allocated across investments. Nevertheless, this is something that many institutions still do, perhaps because it seems so intuitive, perhaps because they don't believe in the separation of capital and risk.⁹ In any case, the diversification effect says that the sum of the standalone risks contributed by individual positions does not equal the total capital required. For example, consider offsetting swap contract positions, which are perfectly negatively correlated. The risk of the combined swaps is zero, so no capital is required to support them. However, the *standalone* capital requirements for each leg would be positive. Clearly the sum of the standalone capital requirements is larger than the capital required for the hedged position; hence the diversification savings.

In addition, the swap legs have equal *negative inframarginal* capital contributions. Removing either side of the hedge would lead to an increase in risk. Hence each leg acts

⁹ See, for example, F. Saita, "Allocation of Risk Capital in Financial Institutions," *Financial Management*, 28, 95-111, 1999.

as hedge against the other. The sum of these negative inframarginal risks is necessarily less than zero; this is just another way to see the diversification benefit. The inframarginals, like the standalones, fail to add up to total capital and therefore can't be allocated.¹⁰ Some take this to mean that there is no practical way to allocate capital sensibly.

However, the *marginal* risks of positions do add up. Think of the combined swap portfolio as pre-existing. Then consider adding a marginal amount of risk to one leg, chosen arbitrarily. Then do the same to the other. The first marginal increase will add positive risk, the second will add equal and offsetting negative risk. Summing the covariances with the pre-existing portfolio results in the right answer: that the marginal risks sum to zero. Thus, as long as the pre-existing portfolio doesn't evolve too quickly, so that we can consider such increments as small, an allocation of capital according to the marginal covariance is internally consistent.¹¹

While the adding up property is helpful, it would nevertheless be a mistake to base capital charges on only marginal capital. The biggest single problem with marginal-capital based charges is that a risk is only costly to the extent it contributes to bank portfolio risk. Systematic risk and intra-corporate agency problems also necessitate capital charges, as discussed below. However, because these do not appear in marginal capital, the resulting allocations cannot alone be used as a basis for pricing.¹²

2. Internal prices must compensate capital providers for systematic risk exposures

This seems to be a non-controversial point, even though many RAROC methods leave out such a charge. The basic idea is that investor required returns on standalone investments include charges for some set of systematic risks. The more systematic risk (i.e., the higher the standard "beta" of an investment) the higher the charge.

As example, consider a simple total return swap on the equity market done at fair value. If this position is uncorrelated with a bank's pre-existing portfolio, it gets no marginal capital allocation at all. The pricing system should evaluate this as a zero-NPV trade, since shareholders could do it themselves on similar terms. However, the expected return on the market is greater than that of cash. So, if the risk system is to view this transaction as zero-NPV (rather than positive NPV), the transaction must be charged the equity risk premium, evaluated on the notional size of the swap. This cost should be levied on the transaction regardless of its relationship to the pre-existing portfolio and regardless of the marginal capital the transaction generates.

¹⁰ See R. Merton and A. Perold, "The Theory of Risk Capital in Financial Firms," *Journal of Applied Corporate Finance*, 6, 16-32, 1993.

¹¹ Froot and Stein (1998a) and S. Myers and J. Read, "Capital Allocation for Insurance Companies," Sloan School MIT, November 2000 make this point formally.

¹² See also Charles Monet, "Using Economic Capital Concepts to Enhance Risk/Return and Manage Capital," JP Morgan, December 2000.

3. Internal prices must compensate capital providers for corporate financing imperfections

This is the portion of pricing marginal capital allocation gets right. Risk charges must increase with the contributed risk of a position. This is a reflection of the scarcity of firm-level capital. Positions should receive a charge that is proportional to the risk they contribute to bank-wide capital. So, for example, if a small amount of the above equity swap contributes twice as much risk as a small credit position, then the equity swap should require twice the amount of marginal capital, and therefore receive twice the capital charge. Note that this is a relative statement only. It describes risk charges *relative* to one another, without saying how large the absolute charge should be.

All that is missing is a single constant of proportionality that applies across positions. Merton and Perold (1993) suggest that the constant of proportionality comes from the deadweight cost of an arm's length option to guarantee the firm against running out of capital. "Deadweight" refers to the difference between the cost of the actual guarantee, and the cost of the guarantee if the firm were totally transparent with no dynamic discretion by managers. Froot and Stein (1998a) suggest that the constant of proportionality comes from the degree of convexity in firm value with respect to capital. The more capital constrained a firm is, the more a given-size reduction in internal funds reduces value. This happens because, with lower internal capital, firm value is more likely to be dissipated through either underinvestment or external financings conducted at punitive prices.

All that is clear from the theory is that the transparent firm with plenty of capital should assign a low proportionality factor and that a firm with little capital should have a high proportionality factor. It is reasonable to think of this factor as a measure of firm-wide risk aversion, although it is difficult to measure risk aversion directly.

Froot and Stein (1998b) summarize these considerations using a two-factor model for pricing marginal risks: the first factor charges according to covariance with the market and the second factor charges according to covariance with the bank-wide portfolio. Idiosyncratic risks that are nevertheless highly volatile are penalized because marginal increases in such risks are correlated with the inframarginal exposures.

4. Internal prices must compensate capital providers for intra-corporate financing imperfections

This is as far as the existing theory goes, but it isn't far enough. Charles Monet, for example, notes that the results of a two-factor approach are inadequate in practice. The covariance with the internal portfolio doesn't rise quickly enough with the size of inframarginal positions. Positions would need to become unacceptably large before receiving appreciable capital charges from the second factor. As position size increases, firms want their capital charges to rise more steeply than portfolio covariance allows.

One reason is that, by itself, the second factor is too simple. It refers to the deadweight costs of replenishing bank-wide capital only. While these costs are legitimate, they shouldn't be interpreted too narrowly: there are also costs to replenishing intra-corporate budgets. Just as firms are punished when they unexpectedly come to the market to replenish capital, so too are intrafirm businesses punished when they unexpectedly come to headquarters to replenish their capital. A number of issues are at work to make such replenishment costly: headquarters worries about the alignment of interests with line managers and about line managers' superior information about line prospects.

This suggests that the risk charge based on the covariance of positions with bank-wide portfolios might be augmented with another charge based on the covariance of positions with line (or area) portfolios. Specifically, it makes sense to charge more if a given position makes it more likely a line budget will either have to be replenished or reinforced through reduced line spending. Value maximization at the firm level would involve minimizing these deadweight costs in addition to minimizing the deadweight costs of obtaining external finance from outside markets.

This logic points to a three-factor model for pricing warehousing activities. The first factor contributes to capital charges based on the product of the price and quantity of market risk. The price of market risk can be interpreted as the market equity premium and the quantity of market risk as the covariance between given position and the market.

The second factor contributes to capital charges based on the product of price and quantity of bank-wide portfolio risk. The price of bank-wide risk can be interpreted as bank-wide risk aversion, and the quantity of bank-wide risk as the covariance between a given position and the entire bank-wide portfolio (inclusive of that inframarginal position).

Finally, the third factor contributes to capital charges based on the product of price and quantity of line-specific portfolio risk. The price of line-specific risk can be interpreted as line-specific risk aversion, and the quantity of line-specific risk as the covariance between a given position and the line-specific portfolio (inclusive of any inframarginal position). With this third factor, an increase in position size increases capital charges more rapidly, since the covariance with line-specific portfolio grows with size. Indeed, since line-specific portfolios are considerably smaller than bank-wide portfolios, capital charges increase much more rapidly with position size. In this way, the addition of a third covariance-based factor makes large unhedged positions more costly to carry than would be implied by contribution bank-wide risk alone.

These ideas can be summarized in a single equation that can be used to determine hurdle rates:

$$k_i = r_f + \beta_{i,m}(k_m - r_f) + \beta_{i,b}Z_b + \beta_{i,l}Z_l$$

where k_i is the hurdle rate for the i th risk exposure, r_f is the risk free rate, $\beta_{i,m}$ is the beta of exposure i and the market, $k_m - r_f$ is the market equity premium, $\beta_{i,b}$ is the beta of

exposure i and the bank-wide portfolio, Z_b is the risk tolerance with respect to bank-wide risks, $\beta_{i,l}$ is the beta of exposure i and the line-specific portfolio, l , and Z_l is the risk tolerance with respect to line- l risks.

Among other limitations, the heuristic approach here assumes that risks are distributed normally, so that the betas are a sufficient description of the risk. Of course, in practice nothing is normally distributed.

But such issues should not obscure the basic point of the equation above: that bank required returns on a risk must exceed those that the broader capital market requires by $\beta_{i,b}Z_b + \beta_{i,l}Z_l$. This is instructive, because it reminds us of the inefficiencies that a bank-located warehousing function faces compared with a special purpose vehicle or mutual fund. It reinforces the idea that warehousing should be limited to those illiquid or nontraded exposures that cannot be passed into the markets and/or exposures in which the bank has some offsetting competitive advantage. This view also creates some justification for higher regulatory capital requirements, insofar as regulators can encourage banks to exit warehousing activities that can be more efficiently accomplished outside of bank balance sheets.

B. Pricing Capital With Other, Non-Warehousing Activities

Of course, the extra two factors create a compelling reason for banks to move away from warehousing activities. Increasingly, financial firm P&L's are driven by origination, distribution and other related service activities. How should we think about pricing and capital budgeting once these activities are important?

In fact, little in the analysis above needs to change. Indeed, all of the above analysis goes through for firms that have no financial businesses. Well-run firms regardless of industry have scarce internal buffers, so the second and third factors remain important. The first factor is traditionally used for capital budgeting in non-financial firms, so it remains as well.

However, there is an important underlying change in emphasis. First, in diversified financial firms, pricing decisions more often apply to business lines or units rather than individual contractual positions. After all, business line risks stress capital more than particular balance sheet exposures. The framework remains useful for benchmarking business lines' profits against the risks that they generate for the firm overall. And it is important to have hurdle rates for business lines, both for capital budgeting and acquisition/divestiture decisions.

Second, there is a change in regulatory emphasis as well. Clearly, bank-wide capital requirements have to be evaluated based not only on balance sheet exposures, but on the business risk of the firm. Regulators need to be mindful of the externalities associated with marrying a fee or service business with a bank. I discussed the positive externalities of such a marriage when I added the expected profitability of service activities to VaR.

However, this marriage also brings negative externalities. Standalone origination activities, for example, themselves carry little systemic risk. A nonfinancial firm may become insolvent or illiquid, but that is unlikely to spread to other firms or to affect the financial sector at large. However, once warehousing and other services are combined in a single firm, systemic and firm-liquidity risks can be triggered from outside of the warehousing sector. This is the reason that a movement toward greater corporate separation of warehousing and origination/distribution activities is beneficial. And when such activities remain co-mingled, regulators need to ensure these added risks are taken into account.

III. Conclusion

This paper has stressed that the distinction between the warehousing and non-warehousing activities of a financial firm is important for understand economic and regulatory capital. Much of the concern with capital adequacy today focuses on the computation of risk and the amount of capital that such risk requires. There is a great deal of attention lavished on computational and measurement issues surrounding price and credit risk, and, increasingly, liquidity risk. These risks – all of which accrue from financial firms' warehousing functions – are certainly important, and sensible measurement of them is a necessary component of any economic or regulatory capital system.

However, given the progress being made on this front, it is important to remember that these warehousing functions are by no means all of the risks that financial firm takes on. They increasingly participate in numerous agency businesses, such as various forms of asset management, account services, and advisory services. These businesses add to the risk of the firm. But importantly they may also contribute to the profitability of the firm, and thereby reduce some of the burdens that would otherwise be born by pre-existing capital as a buffer against risk. This paper has tried to show how these non-warehousing activities impact both our assessment of risk and the capital buffer that a prudent financial firm needs.