Introduction

In the last 30 years, financial systems around the world have undergone revolutionary change. People can borrow greater amounts at cheaper rates than ever before, invest in a multitude of instruments catering to every possible profile of risk and return, and share risks with strangers from across the globe. Have these undoubted benefits come at a cost? How concerned should central bankers and financial system supervisors be, and what can they do about it? These are the issues examined in this paper.

Consider the main forces that have been at work in altering the financial landscape. Technical change has reduced the cost of communication and computation, as well as the cost of acquiring, processing, and storing information. One very important aspect of technical change has been academic research and commercial development. Techniques ranging from financial engineering to portfolio optimization, from securitization to credit scoring, are now widely used. Deregulation has removed artificial barriers preventing entry, or competition between products, institutions, markets, and jurisdictions. Finally, the process of institutional change has created new entities within the financial sector
such as private equity firms and hedge funds, as well as new political, legal, and regulatory arrangements.

These changes have altered the nature of the typical transaction in the financial sector, making it more arm’s length and allowing broader participation. Financial markets have expanded and become deeper. The broad participation has allowed risks to be more widely spread throughout the economy.

While this phenomenon has been termed “disintermediation” because it involves moving away from traditional bank-centered ties, the term is a misnomer. Though in a number of industrialized countries individuals do not deposit a significant portion of their savings directly in banks any more, they invest indirectly in the market via mutual funds, insurance companies, and pension funds, and indirectly in firms via (indirect) investments in venture capital funds, hedge funds, and other forms of private equity. The managers of these financial institutions, whom I shall call “investment managers” have displaced banks and “reintermediated” themselves between individuals and markets.

What about banks themselves? While banks can now sell much of the risk associated with the “plain-vanilla” transactions they originate, such as mortgages, off their balance sheets, they have to retain a portion, typically the first losses. Moreover, they now focus far more on transactions where they have a comparative advantage, typically transactions where explicit contracts are hard to specify or where the consequences need to be hedged by trading in the market. In short, as the plain-vanilla transaction becomes more liquid and amenable to being transacted in the market, banks are moving on to more illiquid transactions. Competition forces them to flirt continuously with the limits of illiquidity.

The expansion in the variety of intermediaries and financial transactions has major benefits, including reducing the transactions costs of investing, expanding access to capital, allowing more diverse opinions to be expressed in the marketplace, and allowing better risk
sharing. However, it has potential downsides, which I will explore in this paper. This focus is not meant to minimize the enormous upsides that have been explored elsewhere (see, for example, Rajan and Zingales, 2003, or Shiller, 2003), or to suggest a reversion to the days of bank-dominated systems with limited competition, risk sharing, and choice. Instead, it is to draw attention to a potential source of concern and explore ways the system can be made to work better.

My main concern has to do with incentives. Any form of intermediation introduces a layer of management between the investor and the investment. A key question is how aligned are the incentives of managers with investors, and what distortions are created by misalignment? I will argue in this paper that the changes in the financial sector have altered managerial incentives, which in turn have altered the nature of risks undertaken by the system, with some potential for distortions.

In the 1950s and 1960s, banks dominated financial systems. Bank managers were paid a largely fixed salary. Given that regulation kept competition muted, there was no need for shareholders to offer managers strong performance incentives (and such incentives may even have been detrimental, as it would have tempted bank managers to reach out for risk). The main check on bank managers making bad investment decisions was the bank’s fragile capital structure (and possibly supervisors). If bank management displayed incompetence or knavery, depositors would get jittery and possibly run. The threat of this extreme penalty, coupled with the limited upside from salaries that were not buoyed by stock or options compensation, combined to make bankers extremely conservative. This served depositors well since their capital was safe, while shareholders, who enjoyed a steady rent because of the limited competition, were also happy. Of course, depositors and borrowers had little choice, so the whole system was very inefficient.

In the new, deregulated, competitive environment, investment managers cannot be provided the same staid incentives as bank
managers of yore. Because they have to have the incentive to search for good investments, their compensation has to be sensitive to investment returns, especially returns relative to their competitors. Furthermore, new investors are attracted by high returns. Dissatisfied investors can take their money elsewhere, but they do so with substantial inertia. Since compensation is also typically related to assets under management, the movement of investors further modulates the relationship between returns and compensation.

Therefore, the incentive structure of investment managers today differs from the incentive structure of bank managers of the past in two important ways. First, the way compensation relates to returns implies there is typically less downside and more upside from generating investment returns. Managers, therefore, have greater incentive to take risk. Second, their performance relative to other peer managers matters, either because it is directly embedded in their compensation, or because investors exit or enter funds on that basis.

The knowledge that managers are being evaluated against others can induce superior performance, but also a variety of perverse behavior.

One is the incentive to take risk that is concealed from investors—since risk and return are related, the manager then looks as if he outperforms peers given the risk he takes. Typically, the kinds of risks that can be concealed most easily, given the requirement of periodic reporting, are risks that generate severe adverse consequences with small probability but, in return, offer generous compensation the rest of the time. These risks are known as tail risks.

A second form of perverse behavior is the incentive to herd with other investment managers on investment choices because herding provides insurance the manager will not underperform his peers. Herd behavior can move asset prices away from fundamentals.

Both behaviors can reinforce each other during an asset price boom, when investment managers are willing to bear the low-probability tail
risk that asset prices will revert to fundamentals abruptly, and the knowledge that many of their peers are herding on this risk gives them comfort that they will not underperform significantly if boom turns to bust. An environment of low interest rates following a period of high rates is particularly problematic, for not only does the incentive of some participants to “search for yield” go up, but also asset prices are given the initial impetus, which can lead to an upward spiral, creating the conditions for a sharp and messy realignment.

Will banks add to this behavior or restrain it? The compensation of bank managers, while not so tightly tied to returns, has not remained uninfluenced by competitive pressures. Banks make returns both by originating risks and by bearing them. As plain-vanilla risks can be moved off bank balance sheets into the balance sheets of investment managers, banks have an incentive to originate more of them. Thus, they will tend to feed rather than restrain the appetite for risk. However, banks cannot sell all risks. They often have to bear the most complicated and volatile portion of the risks they originate, so even though some risk has been moved off bank balance sheets, balance sheets have been reloaded with fresh, more complicated risks. In fact, the data suggest that despite a deepening of financial markets, banks may not be any safer than in the past. Moreover, the risk they now bear is a small (though perhaps the most volatile) tip of an iceberg of risk they have created.

But perhaps the most important concern is whether banks will be able to provide liquidity to financial markets so that if the tail risk does materialize, financial positions can be unwound and losses allocated so that the consequences to the real economy are minimized. Past episodes indicate that banks have played this role successfully. However, there is no assurance they will continue to be able to play the role. In particular, banks have been able to provide liquidity in the past, in part because their sound balance sheets have allowed them to attract the available spare liquidity in the market. However, banks today also require liquid markets to hedge some of the risks associated with complicated products they have created, or guarantees they have offered. Their greater
reliance on market liquidity can make their balance sheets more suspect in times of crisis, making them less able to provide the liquidity assurance that they have provided in the past.

Taken together, these trends suggest that even though there are far more participants today able to absorb risk, the financial risks that are being created by the system are indeed greater. And even though there should theoretically be a diversity of opinion and actions by participants, and a greater capacity to absorb the risk, competition and compensation may induce more correlation in behavior than desirable. While it is hard to be categorical about anything as complex as the modern financial system, it is possible these developments may create more financial-sector-induced procyclicality than the past. They also may create a greater (albeit still small) probability of a catastrophic meltdown.

What can policymakers do? While all interventions can create their own unforeseen consequences, these risks have to be weighed against the costs of doing nothing and hoping that somehow markets will deal with these concerns. I offer some reasons why markets may not get it right, though, of course, there should be no presumption that regulators will. More study is clearly needed to estimate the magnitude of the concerns raised in this paper. If we want to avoid large adverse consequences, even when they are small probability, we might want to take precautions, especially if conclusive analysis is likely to take a long time.

At the very least, the concerns I raise imply monetary policy should be informed by the effect it has on incentives, and the potential for greater procyclicality of the system. Also, bank credit and other monetary indicators may no longer be sufficient statistics for the quantity of finance-fueled activity. I discuss some implications for the conduct of monetary policy.

Equally important in addressing perverse behavior are prudential norms. The prudential net may have to be cast wider than simply...
around commercial or investment banks. Furthermore, while I think capital regulation or disclosure can help in some circumstances, they may not be the best instruments to deal with the concerns I raise. In particular, while disclosure is useful when financial positions are simple and static, it is less useful when positions are complex and dynamic. Ultimately, however, if problems stem from distorted incentives, the least interventionist solution might involve aligning incentives. Investors typically force a lengthening of horizons of their managers by requiring them to invest some fraction of their personal wealth in the assets they manage. Some similar market-friendly way of ensuring personal capital is at stake could be contemplated, and I discuss the pros and cons of some approaches to incentive alignment.

The rest of this paper is as follows. In the second section, I start by describing the forces that have driven the changes. In the third section, I discuss how financial transactions have been changed, and in the fourth section, how this may have changed the nature of financial risk taking. In the fifth section, I discuss potential policy responses, and then I conclude.

The forces driving change

Technology

Technology has altered many aspects of financial transactions. In the area of lending, for instance, information on firms and individuals from a variety of centralized sources—such as Dun and Bradstreet—is now widely available. The increased availability of reliable, timely information has allowed loan officers to cut down on their own monitoring. While, undoubtedly, some soft information that is hard to collect and communicate—direct judgments of character, for example—is no longer captured when the loan officer ceases to make regular visits to the firm, it may be more than compensated by the sheer volume and timeliness of hard information that is now available. Moreover, because it is hard information—past credit record, accounting data, etc.—the information now can be automatically processed,
eliminating many tedious and costly transactions. Technology has therefore allowed more arm’s length finance and therefore expanded overall access to finance.

Such methods undoubtedly increase the productivity of lending, reduce costs, and thus expand access and competition. Petersen and Rajan (2002) find that the distance between lenders and borrowers has increased over time in the United States, and the extent to which this phenomenon occurs in a region is explained by an increase in the bank-loan-to-bank-employee ratio in that region, a crude proxy for the increase in productivity as a result of automation.

Deregulation and institutional change

Technology has spurred deregulation and competition. In the 1970s, the United States had anticompetitive state banking laws. Some states did not allow banks to open more than one branch. Many states also debarred out-of-state banks from opening branches. Banks were small, risky, and inefficient. The reason, quite simply, for these laws was to ensure that competition between banks was limited so that existing in-state banks could remain profitable and fill state coffers.

As information technology improved the ability of banks to lend and borrow from customers at a distance, however, competition from out-of-state financial institutions increased, even though they had no in-state branches. Local politicians could not stamp out this competition since they had no jurisdiction over it. Rather than seeing their small, inefficient, local champions being overwhelmed by outsiders, they eliminated the regulations limiting branching (see Kroszner and Strahan, 1999).

Thus, technology helped spur deregulation, which in turn created a larger market in which technologies could be utilized, creating further technological advances. Both forces have come together to spur institutional change. For example, not only has there been an enormous amount of bank consolidation, but also the activities of
large banks have undergone change. As deregulation has increased competition for the best borrowers, and shaved margins from offering plain-vanilla products to these customers, large banks have reached out to nontraditional customers, or to traditional customers with innovative products.

Taken together, all these changes have had beneficial, real effects, increasing lending, entrepreneurship, and growth rates of GDP, while reducing costs of financial transactions (see Jayaratne and Strahan, 1996, 1998, and Black and Strahan, 2001). Such developments can be seen throughout the world. Let me now turn to how they have changed the nature of interaction in the financial sector and, in the third section, how they may have altered the nature of risks.

**How financial transactions have changed**

*Arm's length transactions or disintermediation*

A number of financial transactions have moved from being embedded in a long-term relationship between a client and a financial institution to being conducted at arm's length in a market. In many parts of the world where banking has been the mainstay, arm's length corporate bond markets and equity markets have expanded relative to the more stable private credit markets. While long-term relationships do lead to greater understanding and trust between parties, they do constrain each party’s choices. Increasingly, only the most complicated, innovative, or risky financial transactions are embedded in relationships—I will have more to say on this shortly.

Greater availability of public information (not just about the client but also about the outcome of the transaction and the behavior of each party), the standardization of financial contracts, and the ability of financial institutions to carve up streams of cash flows (both contingent and actual) into desirable portions have contributed to this process of “commodification” of financial transactions. Consider each of these.
The publicly available credit history of a potential borrower not only expands the set of potential lenders who can screen the borrower, but also serves as a punishment for those borrowers who default by significantly raising the cost and limiting access to future credit. Credit histories are now collateral. Of course, public information does not constrain just borrowers, it also constrains lenders. Large financial institutions dealing with the public are closely scrutinized by the press. They cannot afford to be tainted by unsavory practices. In turn, this knowledge gives retail customers the confidence to enter freely into transactions with these financial institutions.

The standardization of contractual terms allows a loan to be packaged with other contracts and sold as a diversified bundle to passive investors who do not have origination capability. Alternatively, the cash flows from the bundle can be carved up or “tranched” into different securities, differing in liquidity, maturity, contingency, and risk, each of which appeals to a particular clientele. This process of “securitization” allows for specialization in financial markets—those who have specific capabilities in originating financial transactions can be different from those who ultimately hold the risk. Securitization, thus, allows the use of both the skills and the risk-bearing capacity of the economy to the fullest extent possible.

While the collection of data on the growth of the credit derivatives and credit default swaps in the last several years is still in early stages and probably underestimates their usage, the takeoff of this market is a testament to how financial innovation has been used to spread traditional risks (see Chart 1).

Integration of markets

The growth of arm’s length transactions, as well as the attendant fall in regulatory barriers to the flow of capital across markets, has led to greater integration between markets. As Chart 2 suggests, the gross external assets held by countries (claims of citizens on foreigners) has grown seven-fold over the last three decades.
Chart 1
Credit Derivatives and Credit Default Swaps¹
(In Percent of Private Sector Bank Credit²)

¹Credit derivatives from British Bankers’ Association credit derivatives reports.
Credit default swaps from International Swaps and Derivatives Association Market Surveys.
²Includes IFS data on deposit money banks and—where available—other banking institutions
for Australia, Canada, the euro area, Japan, the United Kingdom, and the United States.

Chart 2
External Growth Assets
(In Percent of World GDP)

Source: Lane and Milesi-Ferretti (2005) and IMF staff estimates
The advantages of interlinked markets are many. With pools of capital from all over the world becoming available, transactions no longer depend as much on the availability of local liquidity but on global liquidity. A world interest rate is now close to a reality, with capital flowing to where returns seem the most attractive. In a seminal paper in 1980, Feldstein and Horioka pointed out that there seemed to be a much closer correlation between a country’s savings and its investment than might be suggested by the existence of global capital markets—national investment seemed to be constrained by national savings. The correlation between savings and investment rates within each region has fallen off, dropping from an average of 0.6 in the period 1970-1996 to 0.4 in the period 1997-2004 (see IMF, 2005b, *World Economic Outlook*, forthcoming, fall 2005).

**Reintermediation**

That more financial transactions are conducted at arm’s length does not mean that intermediaries will disappear. For one, intermediation can reduce the costs of investing for the client, even if the relationship between the client and the investment manager is purely arm’s length. “Reintermediation” is given further impetus as the sheer complexity of financial instruments and the volume of information about them increases—investors prefer delegating to a specialist. Transparent institutions, such as mutual funds or pension funds, save transactions costs for investors. Less-transparent institutions, such as venture capital funds or hedge funds, have emerged to search for returns in newer, more exotic areas, as excess returns in more traditional investments have been competed away. Thus, for example, even as equity markets have grown, the share of direct investment by households in markets has fallen off in the United States (see Chart 3).

**Banking relationships in origination, product customization, and innovation**

As more and more financial products migrate to markets, and more transactions are undertaken at arm’s length, are commercial banks (and
their increasingly close cousins, investment banks) becoming redundant? To understand the role banks play, we need to understand the special nature of their capital structure and the relationships they build.

**The role of banks**

Traditionally, a bank has been defined in terms of its twin functions—lending to difficult credits and offering demand deposits, or more generally, payment services. Yet these functions seem contradictory. Why offer depositors liquidity on demand when assets are tied up in illiquid bank loans? Does narrow banking not make more sense, where money market funds invested in liquid securities offer demand deposits while finance companies funded through long-term liabilities make loans? Calls for “breaking up the bank” resurface every few years (see, for example, Simons, 1948, and Bryan, 1988).

Yet the form of the banking organization has remained virtually unchanged over a thousand years, suggesting some rationale for the organizational form. Diamond and Rajan (2001a) argue that it is the
credibility obtained from the fragile capital structure that allows the bank to take on the risks associated with illiquid loans. If the bank mismanages funds, it knows it will be shut down in a trice by its depositors and counterparts in the money and inter-bank markets. It has a very strong incentive to be careful. Since this is widely known and understood, the bank will be trusted by the money market, other banks, and depositors. Its continued access to liquidity then enables it to provide it on demand to those who desire it.5

Risk transfer

Abstracting further to make this discussion more relevant to an industrial economy, the purpose of the bank is to warehouse risks that only it can manage, while financing with a capital structure that gives its management credibility. This means that if some risks become more vanilla and capable of being offloaded to the rest of the financial sector, the banking system will offload them and replace them with more complicated risks, which pay more and better utilize its distinct warehousing capabilities. After all, investment managers, who have a relatively focused and transparent investment strategy, have a lower cost of capital in financing liquid assets and plain-vanilla risks than banks, whose strategies and balance sheets are more opaque (see Myers and Rajan, 1998).6

Consider an example. A fixed rate bank loan to a large corporate client has a number of embedded risks, such as the risk that interest rates will rise, reducing the present value of future repayments and the risk that the client firm will default. There is no reason the bank should hold on to interest rate risk. Why not offload it to an insurance company or a pension fund that is looking for fixed income flows? Increasingly, default risk is also being transferred.7 However, the bank may, want to hold on to some of the default risk, both to signal the quality of the risk to potential buyers, and to signal it will continue monitoring the firm, coaxing it to reduce default risk. The lower the credit quality of the firm, the stronger the role of the bank in monitoring and controlling default risk, as also the greater the need
to signal to buyers. Hence, the size of the first-loss position the bank retains is likely to increase as the credit quality of the loan falls (see Franke and Krahnen, 2005, for evidence).

Thus, risk transfer, through loan and default risk sales, does not completely eliminate risk from bank balance sheets. In fact, bank earnings variability in the United States has not fallen (see Chart 4), and average bank distance to default in a number of countries has not increased (see Chart 5). It is apparent that banks have not become safer despite the development of financial markets and despite being better capitalized than in the past. In fact, they may have well become riskier in some countries. Finally, if we think bank earnings are likely to grow at the rate at which market earnings will grow over the foreseeable future, the declining price-earnings ratio of banks in the United States relative to the market suggest that the market is discounting bank earnings with an increasing risk premium (see Chart 6). This again suggests bank earnings have not become less risky.

Instead of reducing bank risk, risk transfer allows the bank to concentrate on risks so that it has a comparative advantage in managing, making optimal use of its capital while hiving off the rest to those who have a natural appetite for it or to those with balance sheets large enough or transparent enough to absorb those risks passively. It also implies that the risk held on the balance sheet is only the tip of an iceberg of risk that is being created.

Innovation and customization

Apart from originating traditional products, banks also have a role in creating new products. The range of financial needs far exceed the range of financial products that are traded on exchanges. Customized over-the-counter products cannot always be created simply by mixing and matching existing exchange-traded instruments. Instead, banks have to create products tailored to specific client needs.
Chart 4
S&P 1500 Banks: Earnings Volatility
Sample Average of Estimated AR(1)-Process Residuals

Notes: The residual is obtained from regressing annual bank earnings against lagged earnings. In the top panel, each residual is normalized by dividing by the average for that bank across the entire time frame then averaged across banks in the same period. In the bottom panel, a rolling standard deviation of the residuals is computed for each bank and then averaged across banks.
Chart 5
Bank Distance to Default and Trend Component

Source: Datastream and IMF staff estimates
If there is sufficient demand on both sides for a customized product, it may make sense to eventually let it trade on an exchange. Before that, however, glitches have to be ironed out. New financial contracts will not be immediately accepted in the market because the uncertainties surrounding their functioning cannot be resolved by arm’s length participants, who neither have money nor goodwill to spare. For instance, a key uncertainty for a credit default swap is what determines the event of default. Is it sufficient that the borrower miss a payment? Will a late payment on an electricity bill or a refusal to pay a supplier because of a dispute over quality suffice to trigger default? Will a negotiated out-of-court rescheduling of debt constitute default? These are the kinds of issues that are best settled through experience.

If a bank offers the contract to large clients with whom it has a relationship, the unforeseen contingencies that arise can be dealt with amicably in an environment where both parties to the contract are willing to compromise because they value the relationship (this is not to say the occasional dispute will not end in court). Only when contractual features have been modified to address most contingencies can consideration be given to trading the contract on an exchange. Thus, banks are
critical to the process of customization and financial innovation, using their relationships and reputations to test-drive new contracts.

Sometimes the ambiguities in contracts can never be resolved, so the contracts do not migrate to the markets. Take, for instance, a loan commitment—that is, a contract through which a bank agrees to lend at a pre-specified rate if the client demands a loan. Many loan commitments have an escape clause, termed the “material adverse change” clause. This allows the institution to duck the commitment if there is a material adverse change in the client’s condition, a feature that protects the bank from having to make loans in circumstances where they clearly would not be repaid. In turn, this allows the bank to offer cheaper loan commitments. Of course, the loan commitment would mean little if the bank could renege with impunity. Every time an institution invokes the clause without adequate cause, however, its reputation will suffer a bit, and its future commitments will be worth less. This gives it the incentive to invoke the material adverse change clause only in the most necessary circumstances, and the credibility to offer a plausible commitment. Banks, unlike markets, can offer “incomplete” contracts (see Boot, Greenbaum, and Thakor, 1993, and Rajan, 1998).

Finally, there are contracts for which there is only demand on one side. In such cases, banks may be willing to create the necessary contracts, offer them to clients, and hedge the ensuing risks, often through dynamic trading strategies in financial markets.8

This last point suggests that in addition to its traditional role in offering liquidity to clients and the market, banks now also rely on the liquidity of all sorts of other markets to keep themselves fully hedged. We will return to the risks this poses later.

Summary

Let me summarize. Technical change, regulatory change, and institutional change have combined to make arm’s length transactions more feasible. More transactions are now done on markets, as well as by institutions that have an arm’s length relationship with their clients.
This has not, however, marginalized traditional institutions like banks and their relationships. The changes have allowed such institutions to focus on their core business of customization and financial innovation, as well as risk management. As a consequence, the risks borne by traditional institutions have not become any lower. However, now new risks are spread more widely in the economy, and traditionally excluded groups have benefited.

**Are financial systems safer?**

I have outlined a number of changes to the nature of financial transactions. While these have created undoubted benefits and on net are likely to have made us significantly better off, they have opened up new vulnerabilities, to which I now turn.

Let me start by pointing out some vulnerabilities created by the greater reliance of economies on arm’s length transactions and markets. I then will turn to changes in incentives of financial sector managers, which will be my main focus.

**Greater demand on markets**

Markets have become more integrated, have drawn in a greater variety of participants, and, as a result, usually have more depth. Yet the demands made of these markets are not static and typically increase over time.

One reason is that, with the exception of one-time spot deals, arm’s length transactions rely enormously on the superstructure of the market—on trustworthy and timely dissemination of public information, on reliable performance by counterparties (failing which parties expect rapid and just enforcement), on the smooth functioning of the payments and settlements system, and on the availability of reasonable exit options when needed—that is, the availability of liquidity.
The expectation of a reliable superstructure draws participants who are not necessarily financially sophisticated or aware of local nuances (not just the proverbial Belgian dentist but also the return-hungry foreign fund comes to mind). For these investors, continued reliability is extremely important since they do not have recourse to other means of ensuring the security of their transactions.9

Most markets can provide reliability some of the time to all participants and all of the time to some participants. Few can provide it all of the time to all of the participants. So, critical to the resilience of these markets is whether, at times when universal reliability cannot be assured, those who have assurance of reliability can substitute for those that do not. For example, can domestic financial institutions that have a greater ability to manage without the superstructure underpinning markets and that have their own sources of information and enforcement substitute for potentially more dependent foreign retail investors or funds?10

The very forces that broaden access to the markets unfortunately may inhibit such substitution. First, growing perceptions of reliability, accentuated by good times, which tend to paper over all shortcomings, can draw in significant numbers of unsophisticated investors. The tolerance of these investors for ambiguity or for any counterparties who are “nonconforming” may be very limited. As a result, these investors may take fright at the first sign the superstructure or counterparties are under stress, increasing the volume of transactions that have to be substituted for in such times.

Second, the supply of those who can substitute also may fall as a market builds a record of reliability. Knowing that the unsophisticated focus on certain pieces of public information, and that they tend to move markets in ways that are hard to counteract, the sophisticated may reduce their search for alternative, less-public sources of information. The market may become informationally less diverse as it becomes more arm’s length, increasing risks if public information becomes less reliable (in actuality or perception).11 In other words,
while in a “Hayekian” market, aggregating all manner of information is the ideal of market proponents, the incentives for information acquisition may become muted and, instead, market participants may focus excessively on some readily available sources that they believe everyone else is focusing on (also see Allen, Morris, and Shin, 2004).

Equally worrisome, the traditional skills of the sophisticated in managing without a reliable superstructure may fall into disuse. When the accounts of all companies are suspect as a matter of course, each financial institution has plenty of forensic accountants who can untangle the good firms from the bad. As confidence in accounts increases, however, the forensic accountants are let go, leaving institutions less capable of discrimination between firms when corporate scandals emerge.

Put differently, the longer a market’s superstructure proves to be reliant, the more reliance will be placed on it. If it does not improve its systems constantly, it could find that the demands for reliability that are placed on it exceeds its capability of supplying them. The consequence is greater fragility to errors, to misinformation, and to simple bad luck.

**Incentives leading to riskier markets**

Let me now turn to incentives. In my opinion, a potentially greater concern than the market’s superstructure being unreliable is that the managers of the new intermediaries, as well as managers of today’s banks have vastly different incentive structures than bank managers of the past. This is not a bad thing in and of itself. I will argue, though, that these structures could well create perverse incentives in certain situations, and those should be a source of concern.

As I argued earlier, investors have departed banks only to delegate management of their financial investments to a new set of investment managers. Delegation, however, creates a new problem, that of providing incentives to the investment manager. Investors can reward
managers based on the total returns they generate. However, managers always can produce returns by taking on more risk, so investors have to ensure managers do not game them. One common theoretical measure of performance is Jensen’s alpha, that is, the excess returns produced by the manager over the risk-free rate, per unit of risk taken. A sensible way of implementing a performance system based on alpha is to constrain the investment manager to investments in a particular category or style, and evaluate him based on how he performs relative either to others who follow the style or to an appropriate benchmark portfolio with a similar level of risk. In short, the most practical method of providing incentives to managers is to compare their return performance relative to other competing managers who follow broadly similar investment strategies.

Furthermore, the market provides its own incentives. Given that there are economies of scale in investment management (at least up to a point), it makes sense for managerial compensation to be positively related to assets under management, and it typically is. And assets under management are determined by return performance. Even though there is little systematic evidence that past performance by investment managers ensures future performance, investors do chase after managers who generate high returns because they think (incorrectly) the managers have “hot hands.” And current investors, if dissatisfied, do take their money elsewhere, although they often suffer from inertia in doing so. In Chart 7, I present the flows into an average U.S. mutual fund as a function of the returns it generates (see Chevalier and Ellison, 1997). As the chart suggests, positive excess returns (the amount by which returns exceed the returns on the market) generate substantial inflows, while negative returns generate much milder outflows. In short, inflows are convex in returns.

Thus, an investment manager’s compensation is directly related to the returns he generates, but it is also indirectly related to returns via the quantum of assets he manages, which are also influenced by returns. The superimposition of these two effects leads to a compensation
function that is convex in returns, that is, one that encourages risk taking because the upside is significant, while the downside is limited.¹³

The incentive to take risk is most pronounced for managers of young, small funds, where hot high-return strategies, even those that are sure to collapse eventually, may be preferable to steady strategies. The high-return strategy attracts inflows and enhances compensation in the short run, when the cost of failure in terms of foregone future fees is relatively limited. Eventually, if the fund survives, it will have grown large enough that inflows are no longer as welcome because they make the fund unwieldy. The relative cost of losing the franchise through risky investments then will loom much larger, and the fund will become more conservative. Brown, Goetzmann, and Park (2001) show that the probability of liquidation of hedge funds increases with increasing risk, while Chan and others (2005) find that younger hedge funds tend to get liquidated significantly more often, suggesting they do take on more risk.

The emphasis on relative performance evaluation in compensation creates further perverse incentives. Since additional risks will generally

---

Chart 7
U.S. Mutual Funds’ Returns and Net Flows¹

Source: Chevalier and Ellison (1997)
¹Data for young funds (age 2 years).
imply higher returns, managers may take risks that are typically not in their comparison benchmark (and hidden from investors) so as to generate the higher returns to distinguish themselves. While choosing the more observable investments within the benchmark, however, managers typically will be wary of being too different from their peers, for they insure themselves against relative underperformance when they herd. Let us examine these behaviors in greater detail.

**Hidden tail risk**

Consider the incentive to take on risk that is not in the benchmark and is not observable to investors. A number of insurance companies and pension funds have entered the credit derivatives market to sell guarantees against a company defaulting. Essentially, these investment managers collect premia in ordinary times from people buying the guarantees. With very small probability, however, the company will default, forcing the guarantor to pay out a large amount. The investment managers are, thus, selling disaster insurance or, equivalently, taking on “peso” or tail risks, which produce a positive return most of the time as compensation for a rare very negative return. These strategies have the appearance of producing very high alphas (high returns for low risk), so managers have an incentive to load up on them. Every once in a while, however, they will blow up. Since true performance can be estimated only over a long period, far exceeding the horizon set by the average manager’s incentives, managers will take these risks if they can.

One example of this behavior was observed in 1994, when a number of money market mutual funds in the United States came close to “breaking the buck” (going below a net asset value of $1, which is virtually unthinkable for an ostensibly riskless fund). Some money market funds had to be bailed out by their parent companies. The reason they came so close to disaster was because they had been employing risky derivatives strategies in order to goose up returns, and these strategies came unstuck in the tail event caused by the Federal Reserve raising interest rates quickly.
If firms today implicitly are selling various kinds of default insurance to goose up returns, what happens if catastrophe strikes? Will they start defaulting on obligations to policyholders and pensioners precisely when such protection is most needed? It may well be that the managers of these firms have figured out the correlations between the various instruments they hold and believe they are hedged. Yet as Chan and others (2005) point out, the lessons of summer 1998 following the default on Russian government debt is that correlations that are zero or negative in normal times can turn overnight to one—a phenomenon they term “phase lock-in.” A hedged position can become unhedged at the worst times, inflicting substantial losses on those who mistakenly believe they are protected.

Herding

Consider the second distortion: herding. Established fund managers who are evaluated against a common benchmark like the S&P 500 index have an incentive to buy the stocks included in the index as a form of insurance since only severe underperformance triggers dismissal.\(^{17}\) Even if they suspect the stocks are overvalued, they know they will be excused if they perform very poorly when their benchmark also performs poorly.

Would a few enterprising managers not want to buck the trend and, thus, return prices to fundamentals? Unfortunately, few would want to go up against the enormous mass of managers pursuing the trend. The reason is that their horizon is limited. If the mispricing in stocks does not correct itself in a relatively short while, the investment manager will see an erosion of his customers as he underperforms. It takes a very brave investment manager with infinitely patient investors to fight the trend, even if the trend is a deviation from fundamental value. Increasingly, finance academics are coming to the conclusion that prolonged deviations from fundamental value are possible because relatively few resources will be deployed to fight the herd (see, for example, Shleifer and Vishny, 1997, or Lamont and Thaler, 2001).
To summarize, overall incentives to take risk have increased. In addition, however, incentives to take tail risk, as well as incentives to herd and move prices away from fundamentals, have increased. Different managers may suffer from each of these distortions to a different extent. The young and unproven are likely to take more tail risk, while the established are likely to herd more. The two distortions are, however, a volatile combination. If herd behavior moves asset prices away from fundamentals, the likelihood of large realignments—precisely the kind that trigger tail losses—increases. One last ingredient can make the cocktail particularly volatile, and that is low interest rates after a period of high rates, either because of financial liberalization or because of extremely accommodative monetary policy.

Low interest rates and incentives

Low interest rates induce an additional degree of procyclical risk taking into financial markets. Let me illustrate with some examples.

Example 1: Insurance companies may have entered into fixed rate commitments. When interest rates fall, they may have no alternative but to seek out riskier investments. If they stay with low return but safe investments, they are likely to default for sure on their commitments, while if they take riskier but higher return investments, they have some chance of survival. This phenomenon, known as risk shifting (see, for example, Jensen and Meckling, 1976), tends to induce participants to ignore collective downside risks (including illiquidity) since their attention is focused on the upside, the only circumstances under which they survive. Of course, if risk-free interest rates start moving back up, insurance companies can meet their obligations without taking undue risk. Thus, they have an incentive to search for risk when interest rates are low, and to become more conservative when they are high.

Example 2: A second form of induced “risk shifting” can be seen in hedge funds. The typical compensation contract for a hedge fund manager is 1 percent of assets under management plus 20 percent of
annual returns in excess of a minimum nominal return (often zero). When risk-free returns are high, compensation is high even if the fund takes on little risk, while when risk-free returns are low, the fund may not exceed even the minimum return if it takes little risk. Thus, low rates will increase fund manager incentives to take on risk. Since the cost of borrowing also can be low at such times, fund managers can goose up returns by adding leverage. In doing so, they also add risk.\(^{18}\)

In addition to the incentives of managers changing, the quantity of capital seeking riskier investments also can increase when interest rates are low, only to pull back when interest rates rise. Insurance companies, pension funds, and endowments may look to invest in hedge funds so as to increase returns. Young hedge funds are likely to attract significant flows, not just because they are more open to them, but also because everyone knows they will take on additional risk.

Simple proxies such as the VIX for the risk aversion of financial markets in the United States do seem to be positively correlated with the level of short-term interest rates.\(^{19}\) Also, Kodres and Kashiwasi (Global Financial Security Report 2005, box 2.4), among others, show emerging market spreads fall significantly when industrial country interest rates fall unexpectedly, and when interest rate volatility is low (as it is when interest rates are low). This suggests that risk appetites may well increase as interest rates fall, inducing a degree of procyclicality into the financial sector, over and above other sources of procyclicality such as collateral values (see, for example, Hoshi, Kashyap and Scharfstein, 1993; Kiyotaki and Moore, 1997; or Shin, 2005).

**Example 3:** The natural question, then, is why do recipients accept such “hot” money and finance long-term illiquid projects or consumption with them? Don’t they realize that these investors are fickle and likely to evaporate when interest rates rise?

Emerging markets are perhaps the recipients most likely to be damaged by a “sudden stop” imposed by a movement of investment managers toward lower risk as developed country rates rise. Maybe
these governments do not realize the risk they are taking by financing with such “hot” flows. After all, their credit rating improves when global interest rates are low since debt service is easier at low rates. This, in turn, implies they will pay lower risk premia than normal on their borrowing. Spreads can be further compressed if industrial country investors have the additional desire to take on risk discussed above. Faced with falling risk premia and improving credit ratings, it is tempting for the emerging market politician to assign the credit to his own policies (and if they have not actually changed, to the fact that foreign investors have finally recognized how wonderful the policies have been). Why not use the easy money to enhance growth rates? Thus, it is possible politicians might mistake a cyclical phenomenon for a secular trend and initiate a plethora of long-term projects on that basis, only to be forced to liquidate them when the cycle turns.

However, all politicians have to do is to look around and see that everyone seems to be lifted by the rising tide of easy credit to realize their country is not special. This should inspire caution. But politicians themselves have short horizons defined by elections and asymmetric compensation functions. Perhaps oversimplifying, if they can produce high growth before the elections, they will be reelected, while low growth (or high growth) post-elections is discounted, a phenomenon that leads to the “political business cycle” (see, for example, Alesina and Sachs, 1986). They, therefore, have every incentive to accept hot money even if they fully recognize its true character.

In fact, a number of models of early warning systems for crises show that the probability of an emerging market crisis increases when U.S. interest rates rise (see Kamin, Schindler, and Samuel, 2001, or Kamin-sky and Reinhart, 1999). The point, therefore, is that common factors such as low interest rates—potentially caused by accommodative monetary policy—can engender excessive tolerance for risk on both sides of financial transactions. When markets are integrated, these phenomena can spill across borders. An industrial country’s monetary policy, while appropriate for its domestic circumstances, may be inappropriate taking into account spillover effects on other economies.
Financial risk and real risk

Thus far, we have discussed excessive financial risk taking. When is it most likely to translate into real risk? Clearly, excessive risk tolerance leads to an excessive willingness to finance real investment, with the potential for overcapacity and a waste of real resources to society.

Are there other consequences? Can the buildup of financial claims independently affect real activity? In the perfect world, the answer should be no. If there is an adverse shock to the real value of an asset or position, it should be immediately reflected in the financial claims that are written on that asset or position. Holders of financial claims will absorb the loss, and since the aggregate loss cannot be more than the real loss suffered, the consequences cannot be any greater.

In practice, however, real losses need not be allocated so smoothly to financial claims. First, with a myriad of complex claims written on the same underlying real asset, there may be legal uncertainty on who bears what loss. Bankruptcy procedures, which are an attempt to allocate losses, can be prolonged. During this time, gaming between firm managers and claimants, between various classes of claimants, and between claimants and customers can affect real activity, converting financial distress into economic distress.

Second, in anticipation of a complicated bankruptcy, or when there is no such procedure to restrain claimants, some short-term claimants may prefer to press their financial claims, forcing a liquidation of long-term real assets. Emerging markets have experience with this. A downturn in a country’s economic prospects can lead not just to a sudden stop of new capital but also a forced liquidation of existing real assets as investors run (see Calvo, Izquierdo, and Mejia, 2004; Diamond and Rajan, 2001b; Jeanne, 2002).

The adverse consequences I have just described are multiplied when there is too little financial liquidity in the system. Liquidity
allows holders of financial claims to be patient, allows the netting of offsetting claims, and allows the value of the net financial claim to more fully reflect fundamental real value. Not only does illiquidity perpetuate the overhang of financial claims as well as uncertainty about their final resolution, a perception of too little aggregate liquidity in the system can trigger off additional demands for liquidity (see Diamond and Dybvig, 1983) and create contagion (Diamond and Rajan, 2005). Since systemic instability, rather than the demise of a few players, is the real concern, let us now turn to whether the traditional providers of liquidity will step up to play their role.

**Can the traditional providers of liquidity be relied upon?**

The additional need for liquidity as financial systems have become more arm’s length stems from a variety of reasons. First, as new and inexperienced players are drawn in, there is a greater need for resilience in the system. Second, because more complicated instruments are used and larger positions can be built up, a lot more uncertainty needs to be tided over. Finally, large movements of asset prices away from fundamentals, as well as a concentration on tail risks, implies that in certain eventualities, significant losses will have to be allocated. Can banks step up to provide the needed liquidity?

**Fundamentals and asset prices**

Start by asking whether banks and investment banks will take contrarian positions and reduce the risk associated with misaligned asset values and excessive concentrations of risk. The answer is likely to be no. For one, banks typically do not have the mandate to take the large trading positions necessary to arbitrage away misalignments. Moreover, while bankers may not have the same skewed incentives as investment managers, they also are responsible for quarterly profits. They therefore face the same limits to arbitrage as do investment managers—they cannot carry a losing position for too long.
More problematic, however, is that because they typically can sell much of the risk off their balance sheets, they have an incentive to originate the assets that are in high demand and, thus, feed the frenzy. If it is housing, banks have an incentive to provide whatever mortgages are demanded, even if they are risky “interest-only” mortgages. In the midst of a frenzy, banks are unlikely to maintain much spare risk-bearing capacity. If the returns to originating risk are high, and banks have to keep a piece of every risk originated, they have every incentive to utilize their balance sheets fully when the frenzy is on, rather than buck the trend (and their profitable peers) and keep spare capacity for a potential, low-probability crisis.

This also means that to the extent that the frenzy has contributed to a misalignment of asset prices, bank balance sheets are unlikely to be fully insulated from a realignment of those prices with fundamentals. In addition to explicit bank liabilities, significant downturns can bring along a number of contingent and implicit commitments. Given that banks are increasingly in the guarantee business and are concerned about their future reputations, they will attempt to fulfill these commitments, especially while the extent of the downturn is still unclear. This will further weaken the quality of their balance sheets. Finally, in an arm’s length market, participants always have recourse to litigation, which can bring further uncertainty and liabilities to the balance sheets of those who are perceived to have fed the frenzy.

In summary then, while traditional providers of liquidity may have improved their abilities to manage risk, they are forced by competitive forces to remain on the frontiers of risk creation and absorption. They may not maintain much spare risk-bearing capacity or liquidity-creating capacity. Can they, however, at least reallocate liquidity as they have done in the past?

Reallocating liquidity

One of the advantages of banks is that they can be well-diversified across liquidity needs in the system and can thus provide liquidity
most efficiently to those who need it (see Kashyap and others, 2002). Gatev and Strahan (2004) show that when the commercial paper market dried up for many issuers following the Russian crisis in 1998, banks with higher levels of transaction deposits were perceived (by the stock market) to be lower risk. They experienced larger inflows of funds—possibly from investors fleeing the commercial paper market—just when firms denied access to commercial paper started taking down backup lines of credit from these banks. Thus, banks helped intermediate liquidity back into the system in a time of crisis. As the crisis passed, backup lines were repaid, and deposits flowed back into commercial paper. In a sense, banks provide guarantees of liquidity to borrowers and depositors, but because these came at different times in 1998, they were naturally hedged. Central banks also can help by increasing overall liquidity, as did the Federal Reserve in 1998.

One reason banks functioned effectively in 1998 as liquidity providers of next-to-last resort was because they were not perceived as credit risks, so they attracted spare liquidity. This need not be the case in future crises. As originators of credit risk, banks hold first-loss positions. A downturn in the economy, perhaps caused by and causing asset price realignment, would result in actual losses being borne by bank balance sheets. Of course, banks hedge some positions through dynamic options-like strategies. As liquidity falls and prices move more, they will have to make larger and larger trades to offset those initial losses, thus demanding liquidity from the market precisely when it has little to offer. Bank losses could widen.21

In sum, while the experience in the 1998 crisis in the United States should offer a certain measure of confidence that liquidity will continue to be provided in a crisis, one should not become overly sanguine. If banks also face credit losses and there is uncertainty about where those losses are located, only the very few unimpeachable banks will receive the supply of liquidity fleeing other markets. If these banks also lose confidence in their liquidity-short brethren,
the inter-bank market could freeze up, and one could well have a full-blown financial crisis.

Summary

So, on net, what can we say about how the stability of the financial system has evolved as the nature of the system has changed? While the system now exploits the risk-bearing capacity of the economy better by allocating risks more widely, it also takes on more risks than before. Moreover, the linkages between markets, and between markets and institutions, are now more pronounced. While this helps the system diversify across small shocks, it also exposes the system to large systemic shocks—large shifts in asset prices or changes in aggregate liquidity. The incentive structure of investment managers, as well as intensified competition, may contribute to “endogenizing” the large systemic shocks (see Knight, 2004, for a nice exposition of endogenous risk). Not only might investment managers have a greater tendency to allow asset price misalignments, they also may have a tendency to leave themselves exposed to events in “the tail” of probability distributions, without preparing adequately for them. Tail events may prompt a flight to quality and liquidity. Unfortunately, traditional providers of liquidity could find it harder to step up at such times.

While it is hard to be categorical about anything as complex as the modern financial system, it is possible that these developments are creating more financial-sector-induced procyclicality than in the past. They also may create a greater (albeit still small) probability of a catastrophic meltdown. Unfortunately, we won’t know whether these are, in fact, serious worries until the system has been tested.

It is true the volatility of growth in industrial countries has been falling, partly as a result of the increasing flexibility of real economies, partly as a result of better policies, partly as a result of increased trade, and partly as a result of better financial markets. But the nature of tail risks, especially those related to credit, is such that we should not be lulled into complacency by a long period of calm. The absence of
volatility does not imply the absence of risk, especially when the risk is tail risk, which may take a long time to show up. Moreover, it is not even clear that volatility in the financial sector has fallen. In Chart 8, we plot the volatility of GDP growth in the United States and stock market returns (quarter on quarter returns on the S&P 500 index, estimated over a rolling 12-quarter period). While there seems to be a clear trend down in the volatility of GDP, there does not seem to be such a trend in the stock market.

It is also true the financial system has survived some large shocks in the past, under the able stewardship of Chairman Greenspan. The crash of 1987, the world panic of 1998, the bursting of the stock market bubble in 2000-2001 were all shocks that might have had far worse consequences for industrial country financial sectors, and for the world economy, were it not for appropriate intervention by monetary authorities.

Nevertheless, the experience thus far should not make us overly sanguine. First, can we be confident that the shocks were large enough and in the right places to fully test the system? After all, a shock to equity markets, though large, may have less effect than a shock to credit markets. In other words, can we be confident that all losses are equal, or will some losses be more systemic than others? Perhaps Chairman Greenspan should be faulted for allowing only two mild recessions during his tenure. And perhaps we can sleep better at night if we pray, “Lord, if there be shocks, let them be varied and preferably moderate ones, so we can stress-test our systems.”

Second, some argue it is better to pick up the pieces after the fact through a liquidity infusion by monetary authorities, than attempt to second-guess the market and prevent risk buildup. There seems to be a presumption in this argument that liquidity infusion is costless. It is not. It does impose lower policy rates, sometimes for a considerable duration, and entails a tax on savers and a transfer to those who need the liquidity. The low rates implicit in liquidity intervention also could create their own incentive distortions.
Market-friendly policies to prevent or discourage excessive risk taking should be considered in the light that most of the remedies we currently contemplate are costly.

Before I come to policy, let me turn to an important question. If indeed risk taking is excessive, why don’t investors offer their managers compensation contracts that restrain the short-term emphasis on returns and associated risk taking, as well as encourage them to maintain adequate liquidity? There is a *prima facie* case for policy intervention only if private incentives for risk taking or liquidity provision depart from what is socially desirable. Even if this case can be established, however, intervention will be warranted only if its benefits exceed its costs. But let us start by asking whether there is a *prima facie* case.

*Do investors have the incentive or ability to restrain managers?*

**Restraining risk taking**

Investors indeed may have too little private incentive to restrain the focus of investment managers on generating returns in the short run.
For one, there is very little systematic evidence that past performance in financial investment is an indicator of future performance, Warren Buffet or Peter Lynch notwithstanding. This implies that the constant movement by investors between funds has little social value (and potentially negative value—see Frazzini and Lamont, 2005). But current investors in an individual fund benefit when new investors pour in because the fund’s average costs go down. As a result, the private gains from attracting new investors through a fund’s superior short-term performance exceed the social value, and current investors have too little incentive to restrain managers from focusing on the short term.

Similarly, because issuers—such as emerging market politicians—also may be myopic and may not internalize the cost of accepting hot money fully, investors may not bear the full cost of the real damage inflicted as their investment managers herd in and out of investments.

Even if they wanted to provide their managers appropriate incentives, however, investors may not have the ability to do so. For one, they may not be able to penalize a manager who follows the herd into disaster. After all, it is easy for the manager to walk away and get a job in another fund, blaming the collective crash for his poor performance.

Equally important, however, they may not have complete control over managers because of weaknesses in corporate governance, for example. If so, and because managers have private incentives to generate returns in the short term (to preserve their jobs or for the public adulation that success brings), the private equilibrium again may generate excessive risk taking.

Restraining illiquidity

Liquidity is clearly a public good, so it is likely the private incentives to provide for it, or not overuse it, are inadequate. It is hard for a private actor to fully capture the benefits of providing liquidity; if prices are higher and more closely reflect fundamentals, all who trade
benefit, not just the actor who injected liquidity into the market (Bhattacharya and Gale, 1987; Caballero and Krishnamurthy, 2005). Therefore, the private sector has too little incentive to provision for it.

In fact, it makes sense for an investment manager to free-ride on the liquidity provided by others. Hot money departing an emerging market does not pay for the liquidity it relies upon as it leaves. Thus, both risk taking and liquidity provisions by investment managers may not be adequately policed by private markets (see Chakravorti and Lall, 2004, for a related but different rationale for why markets may not write the right incentive contracts with investment managers).

Given that there is a prima facie case, can we say anything more? Any intervention has costs. Chairman Greenspan, in a speech some years ago, suggested some very useful guidelines for interventions: “Proceed cautiously, facilitate and participate in prudent innovation, allow markets to signal the winners and losers among competing technologies and market structures, and overall—as the medical profession is advised—do no harm.” This would suggest market-friendly, low-cost intervention, but only if we can show the benefits outweigh the costs.

Unfortunately, since we do not know the probability of a potentially catastrophic meltdown of the financial sector (though it is likely to be small), it is hard to do a precise cost-benefit analysis. Greenspan (2004) suggests that a risk management approach to policymaking should take into account small-probability events, especially if the potential costs of those events is likely to be large. Ultimately, though, even a risk management approach boils down to judgments about costs and probabilities, and at present, these will be subjective. More data certainly needs to be collected on the magnitude of these risks.

Nevertheless, I would argue that given the potential costs of the concerns I raise, if we can find low-cost ways of nudging excessive risk taking down, and making it less procyclical, we should use them. Let us examine the tools we might have.
Monetary and regulatory responses

The two main tools that suggest themselves are monetary policy and prudential regulation and supervision. I start with why monetary policy might have a role.

Changes in the process of credit expansion and contraction

One of the main changes under way is that banks increasingly account for a smaller and smaller fraction of expansion in financing. Much more is now fueled by organizations like hedge funds and pension funds. As we have argued, changes in the rate environment can significantly affect the assets under management in risk-tolerant institutions, as well as the allocation of investments by those institutions. This might induce substantial additional procyclicality into monetary policy, as well as increasing the small probability of a serious crisis.

Details matter. For example, we have seen that changes from a high interest rate environment to a low interest rate environment could leave a number of institutions with long-term fixed interest liabilities searching for yield and, hence, risk and willing to expand lending. Also, the change can set off a self-reinforcing upward revaluation of asset prices, collateral values, and credit (see, for example, Bernanke and Gertler, 1990; Kiyotaki and Moore, 1997; and Shin, 2005). Conversely, a change from a low interest rate environment to a high interest rate environment can induce a flight to quality with attendant effects on high-risk illiquid credits, and a collapse in asset price growth.

The level of interest rates also will matter. In a persistently low interest rate environment, more institutions (such as hedge funds) that are compensated on the basis of nominal return will stretch for yield by taking on risk, including maturity risk. Note that via these actions of financial institutions, changes in policy rates will affect
long rates, contributing to any existing channels of monetary transmission from policy rates to long rates.23

**Implications for monetary policy**

The above discussion, however, has the following implications for a risk management approach. First, rapid, large changes in monetary policy have significant costs, not just in the domestic economy but in all interconnected markets (monetary authorities clearly recognize this—see, for example, Ferguson, 2003). Adjustment costs need to be factored into policy. This will typically imply a more measured optimal policy than one that does not factor adjustment costs. Of course, “measured” does not necessarily imply “perfectly predictable.” Some two-way uncertainty is always useful to prevent excessive gaming or complacency by market participants.

Second, while deflation can be immensely harmful for the real economy, an unanticipated but persistent low interest rate can be a source of significant distortions for the financial sector, and thence for asset prices. Not only does this mean staying further away from deflation so that extremely low policy rates do not have to be used as a tool, it also implies exercising greater supervisory vigilance when those rates are in effect to contain asset price bubbles.

Third, and somewhat obviously, one can no longer just examine the state of the banking system and its exposure to credit to reach conclusions about aggregate credit creation, let alone the stability of the financial system. While the banking system still may be the lever by which the entire financial system is controlled, other parts could have substantial influence.

Finally, given that the financial sector may experience greater liquidity and solvency problems in really bad states of the world, central banks have to be vigilant for any possible shortfalls in aggregate liquidity. A risk management approach will imply the central bank will do everything to avert those states by pumping in liquidity.
The danger, of course, is that this could induce moral hazard in market participants. Clearly, any specific assistance will have to include penalties for those managers who have left their institutions overexposed. Central bank credibility in enforcing these penalties will go a long way in limiting moral hazard.

**Prudential supervision**

Some economists argue that monetary policy alone is insufficient to address concerns about both inflation and financial stability, so a second instrument with countercyclical characteristics is needed (see, for example, Goodhart, 2005). A natural candidate seems to be some aspect of prudential supervision.

Two questions arise immediately. Who should come under the prudential supervision net? And what instruments should be used?

**Who should be supervised?**

Large institutions at the core of the financial sector will have to be supervised. However, some might argue that there is no need to bring institutions like hedge funds under the regulatory umbrella. After all, if rich, sophisticated investors want to lose their money by placing it with some mathematical whiz kids who think markets mimic their models, why should regulators intervene?

Clearly, there is no public policy rationale for protecting sophisticated investors. And despite their aura, few hedge funds are large enough to pose systemic risks, long-term capital management (LTCM) notwithstanding. In the jargon, there is no micro-prudential reason to regulate hedge funds. However, there may be a macro-prudential rationale because funds may herd on some trades, causing both large asset price movements as well as demands on liquidity when these trades are reversed, with potential systemic consequences. For this reason, and because accurate asset prices as well as liquidity
are public goods, there is some rationale for bringing such institutions under the supervisory net.

Even if there is such a rationale, do supervisors have the ability to limit excessive risk taking and herding while at the same time not chilling the legitimate and valuable functions such institutions provide? That depends on what instruments are likely to be used and what the effectiveness to intrusiveness ratio is likely to be. Let us turn to that now.

**Instruments: Supervision and disclosure**

For institutions such as small banks that directly hold assets on their balance sheets to maturity, supervisors can monitor positions and slow the creation of assets in hot sectors through market-friendly means. For example, they can require lower loan-to-value ratios after a sustained bout of house price inflation, or caution banks on making interest-only loans at such times.

Direct supervisory oversight of each position or public disclosure is likely to be ineffective for institutions with less homogenous positions and with active trading strategies. Unlike with bank loans, a sample of current positions is likely to say little about an institution’s aggregate risk exposure. Supervisors really need to know all positions as well as the institution’s dynamic trading strategy in order to judge risk exposure effectively, and this will be prohibitively costly (as well as useless soon after, as the institution trades away from the position).

More useful would be for supervisors to gauge the risk management structure of the institution, the risk models the institution uses, as well as to require stress tests/reports of sensitivities of portfolios to changes in macro variables or correlations with other asset prices. Some of this information, stripped of details that would reveal specific positions, could be shared so as to make participants aware of risk concentrations.
While transparency and disclosure are often useful, they are, however, no panaceas. We need to understand better how we can communicate the extent of risk that is being undertaken by an institution in a way that is easy to grasp for both supervisors and investors. We also need to factor in the public or regulatory response to such disclosures, as well as the institution’s counter-response (for an excellent treatment of the perverse response induced by marked-to-market accounting, see Plantin, Sapra, and Shin, 2004). All said and done, we have some way to go before transparency and disclosure offer effective tools for the supervisor and the public to exercise discrimination.

**Instruments: Capital**

Goodhart (2005) suggests that capital requirements should be procyclical. For instance, the capital adequacy requirement on mortgage lending could be related to the rise in house prices (relative to core inflation), while the requirement for lending to construction and property would be related to the rise in property prices.24 This is an ingenious suggestion, which attempts to “lean” against the asset price inflation without stopping it dead in its tracks. It may, however, have wider consequences than only restraining aggregate lending. Institutions always can issue additional capital, and the ones that will find additional capital “cheapest” are those whose asset portfolios are most concentrated in the boom sectors and have, therefore, seen their own stock price run up. Thus, the pace of overall lending could slow, but the lending could become more concentrated in a few institutions that specialize in property. Of course, if loans can be sold to unsupervised institutions, even the pace of overall lending might be only marginally impacted.

The broader point is that capital acts as a budget constraint for the risk a particular institution can take, but is less effective in setting aggregate macro-prudential limits, especially if new capital can be raised or if a number of financial institutions are outside the supervisory umbrella. The way capital requirements are implemented also makes them relatively ineffective against risks that occur in the tails,
which may give institutions an added incentive to load up on these risks. Finally, while capital requirements provide some buffer against insolvency, they provide none against illiquidity.25

Instruments: Incentives

If external restraints like disclosure or capital requirements are likely to be ineffective, perhaps the focus should shift to ensuring investment managers have the right incentives. How does the “market” ensure that investment managers have the right incentives, they are not too myopic in their investment strategies, and they internalize the risks they take?

One way the market induces the right incentives is by requiring managers to have some of their own wealth invested in the funds they manage. For instance, as of December 2004, 811 of the 2,364 managers in the TASS database of hedge fund managers said they had personal investments in their own funds. Only 179, however, reported the quantity of personal holdings. On average, managers had personal investments amounting to 16 percent of fund assets. The median amount was 7 percent; the minimum 0.02 percent.26 Similarly, the usual percentage that partners in venture capital funds are expected to contribute is 1 percent of the committed capital. This number can be lower for very large funds and for buyout funds (which often are very large). Usually it is the senior partners who have to contribute more of this money, while the younger partners invest less.27

A market-friendly way to create better incentives for investment managers is to encourage the spread of this practice. Industry groups could urge all managers to vest some fixed portion of their pay (which could be a norm varying across different segments of the investment management industry) in the funds they manage. A fraction of pay might be better than a fraction of total capital—a norm in which managers start by owning a minimum share of total capitalization might serve as an entry barrier, restricting entry into investment management only to the wealthy. Furthermore, the fixed fraction of
pay has countercyclical properties in that managers who generate high returns (and thus get high pay) will invest more in their funds, which will counteract incentives to take overly risky bets.

In order that incentives be to invest for the long term, the norm could be that the manager’s holdings in the fund would be retained for several years (say a significant part of the duration of a typical asset price cycle). Of course, such holdings will not eliminate the factors that lead a manager to focus on the short term, only offer an offsetting incentive.

This is clearly, again, no panaceas—the managers of LTCM did have substantial stakes in their enterprise. Some other immediate concerns include the following.

Will the manager become overly conservative? Additional conservatism is an objective, which, if carefully moderated, is in the public (though perhaps not the private) interest. However, one should recognize that some institutions provide a public service by taking risks, and the share of compensation, as well as the duration for which assets are held in the fund, will have to be modulated based on the industry segment the investment manager is in.

Will the manager take excessive risks to increase asset values? If the manager holds shares in a highly levered institution, he may have an incentive to increase the level of risk, especially if the institution is likely to default. This is why the norm in which the manager holds the entire range of securities of the institution (or be paid in phantom securities that mimic the asset holdings of the institution) becomes important.

Will the manager manipulate accounting or reduce reserves to boost securities values? This is always a possibility, but the long-term escrowing will serve as a disincentive for managers to engage in short-term securities price manipulation. It is far more difficult to manipulate accounts over the long term.
What happens if the manager quits or is fired? It would be sufficient to hold his securities in escrow for, say, a year after he leaves, so he does not game his departure strategically, and then give them to him to dispose as he chooses.

Will the manager trade out of his position? It is possible that the manager could undertake offsetting trades that essentially hedge him against the risk of the positions held in the fund. Clearly, such violations can be detected (with some difficulty) in the same way as insider trading is detected, provided the agencies of the government support the self-regulatory groups set up by industry. It does lead to questions of whether self-regulation by industry groups should be accompanied by stronger government-supported sanctions than simply expulsion against those who violate industry norms.

More generally, while I have offered one suggestion on how incentives resulting in an excessive focus on generating returns in the short run might be ameliorated, there are obviously others. My intent here is not to propose yet another layer of regulation, but to start a debate on whether a focus on incentives might be more effective than ever-more-onerous disclosure and capital requirements. “Personal” capital requirements might be more effective than traditional firm-level external capital requirements in motivating an investment manager, and deserve consideration as a complement, or even a substitute to existing requirements.

Modalities

Industry groups should see self-regulation of incentives to be in their own interest. After all, everyone is likely to invest in the same assets whose prices may be driven away from fundamentals, everyone relies on the same pool of liquidity, and the collective benefit from the constant churning by investors between funds is small. Self-regulation might be better than public regulation and supervision because norms can be set by those with experience in the business and can be monitored by those who know how the norms can be
evaded. Industry groups also have the incentive to keep the cost of bureaucracy down.

However, one should not underestimate the costs of setting up an apparatus for monitoring and enforcing norms or overestimate the political will to do so. Countries with developed markets and powerful investment industries may not have the political consensus to encourage desirable steps. Could emerging markets, which suffer from the volatility imposed by hot money and are faced with the stark choice of taking the risk associated with hot money or imposing (often ineffective) one-size-fits-all capital controls, press for better incentives? Or will it take the rise of unregulated investment managers in emerging markets for a global consensus to emerge? One can do no more than speculate on the answer.

Summary

I have not discussed concerns about too big to fail, about too complex to fail, etc. These are important issues, but I have little to add to excellent treatments elsewhere (see, for example, Stern and Feldman, 2004).

What is clear is that we are a long way from knowing all the answers on how to reduce the risk of financial instability. It seems obvious that we should use all the innovations that finance has created in order to prevent market forces from creating excessive risk. We have to steer between the Scylla of excessive intervention and the Charybdis of a belief that the markets always will get it right.

Conclusion

Technological change, market liberalization, and institutional change have combined to expand access to credit and risk-sharing opportunities. While many of these changes are most pronounced in the United States, they are making their way to other countries. Furthermore, to the extent that both goods and financial markets are
increasingly interlinked across the world, no country will be immune from the consequences of these changes.

I believe the changes have, in general, expanded opportunities significantly and, even on net, have made the world tremendously better off. But opportunities can be used for good and for bad. This is why it is so critically important to get incentives right. Given the possibility of perverse incentives coming together in some states, a risk management approach to financial regulation will be important to attempt to stave off such states through the judicious operation of monetary policy and through macro-prudential measures. I argue some thought also should be given to attempting to influence incentives of financial institution managers lightly, but directly.

Risk never can be reduced to zero, nor should it be. We should be prepared for the low probability but highly costly downturn. In such an eventuality, it is possible the losses that emanate from a financial catastrophe cannot be entirely borne by current generations and are best shared with future generations. Some of the mechanisms for sharing such systematic risks with future generations, such as (defined benefit) social security, are being changed. While there are gains from doing so, and from ensuring their sustainability, we need to ensure that the intergenerational risk-sharing mechanism they offer is not overly weakened. We also need to continue improving the intrinsic flexibility of our economies, so as to better ride out the downturns that, almost inevitably, will occur.

Author’s note: This paper reflects the author’s views and not necessarily those of the International Monetary Fund, its management, or its board. The author thanks Laura Kodres for extremely useful conversations and suggestions; Sergei Antoshin for valuable research assistance; and Douglas Diamond, Jonathan Fiechter, Laura Kodres, Donald Kohn, Hyun Shin, Jeremy Stein, and Hung Tran for valuable comments on a previous draft.
Endnotes

1In the jargon, investment manager compensation is likely to be convex in returns, while bank manager compensation in the past was more concave. This difference creates a difference in risk preference.

2One might think that the amount of real risk in the economy should limit the amount of financial risk—after all, the underlying cash flows have to be shared among participants, and for every financial instrument that inflicts a loss, there is a counterparty who gains. This is true in a static sense only if financial distress does not cause economic distress. If, however, financial losses cannot be allocated smoothly—for example, because of illiquidity—they may have real consequences. Obviously, in a dynamic sense, greater financial expansion can create greater real risk as, for example, when too many projects are started only for many of them to be shut down prematurely.

3In the jargon, pooling reduces adverse selection—the probability that the seller cherry-picks the bad assets to sell. Since the performance of a pool is more predictable than the performance of any individual asset, the buyer (and the market) can attribute any underperformance of the pool to misbehavior by the seller. Since sellers are typically repeat sellers in this market, this is enough to keep them honest. The tranching of the cash flows from the pool allows the risk of default to be stratified with the first loss borne by the seller. This reduces moral hazard, that is, the risk that the seller will not monitor the risks carefully after selling much of his stake.

4A typical synthetic collateralized debt obligation (CDO) has an equity or first-loss tranche (absorbing approximately the first 3 percent of the losses), a mezzanine tranche (absorbing between 3 percent and 7 percent of the losses), a senior tranche, and possibly a super senior tranche absorbing the rest. Thus, the senior tranche will not suffer a loss unless losses exceed 7 percent of the notional amount of the assets.

5Note that the depositor in this framework could be replaced by anyone who demands liquidity—a firm that demands the bank honor a draw-down on a line of credit tests the bank as much as a depositor who demands his money back. It is in this sense that the bank gains credibility by promising liquidity. In fact, to the extent that the bank attracts both those who demand liquidity and those who supply it, it can offset one against the other, going to the market or its reserves only for the net amount (see Kashyap, Rajan, and Stein, 2002). Gatev and Strahan (2004) find that banks were at the center of liquidity inflows during the 1998 crisis, which enabled them to provide liquidity to stressed firms.
An interesting study in this regard is Morgan (2002), which finds that uncertainty about bank ratings (as reflected in splits between Moody’s and S&P about the rating of a bank’s bonds) has increased markedly since 1986. Uncertainty about banks is also markedly higher than about other industries.

For an extensive analysis of risk transfer, see the *Global Financial Stability Reports* published by the IMF, 2003-2005.

It is worth noting that financial institutions that create financial products to serve a general clientele may have a tendency to become more complex. The least-cost financial product often makes use of multiple markets. For instance, the solution to a client firm’s problem may require the financial institution to sell the firm a customized over-the-counter option, which is then hedged dynamically using other markets. The more markets the financial institution operates in, the easier it will be to come up with the right solution (and the fewer the conflicts of interest in pushing it), and the easier it will be to hedge. Of course, the diseconomies of scope and of organization will eventually kick in, increasing the transactions costs of doing business within the institution. Whether financial institutions have become overly complex is an issue I will not address in this paper.

For example, Qian and Strahan (2005) find that foreign bank’s ownership of a country’s loans is higher when the costs of using the country’s legal system are lower.

Mian (2004), for example, finds that domestic banks in Pakistan are much more likely to restructure troubled loans using a private workout, whereas foreign banks are more prone to use courts.

One example of a focus on easily available public information is the use of corporate bonds ratings in the CDO market, even though such ratings may be misleading when applied to highly leveraged instruments (see BIS, 2005).

See Agarwal and others (2004), Chan and others (2005), and Chevalier and Ellison (1997), for example.

Not all developments take us in the direction of increasing short-term behavior. For example, as the *Global Financial Stability Report* September 2005 points out, insurance companies are increasingly looking for long-term investment strategies and longer-term instruments, so as to hedge their long-term liabilities. My arguments should be seen as reflecting where I think the central tendency is rather than a universal trend.

Of course, much of the exposure of these firms comes from holding senior tranches, where they will suffer loss only if there is a serious downturn in the
economy and the protection afforded by the originator’s equity and the mezzanine levels proves insufficient.

15Peso risk is named after the strategy of investing in Mexican pesos while shorting the U.S. dollar. This produces a steady return amounting to the interest differential between the two countries, although shadowed by the constant catastrophic risk of a devaluation. Another example of a strategy producing such a pattern of returns is to short deep out-of-the-money S&P 500 put options (see Chan, Getmansky, Haas, and Lo, 2005).

16Certainly, the pattern of returns of hedge funds following fixed income arbitrage strategies suggested they were selling disaster insurance. The worst average monthly return between 1990 and 1997 was a loss of 2.58 percent, but losses were 6.45 percent in September 1998 and 6.09 percent in October 1998.


18None of this is to take away from the positive role hedge funds play in searching for the highest yields and in providing liquidity to markets.

19A Generalized autoregressive conditional heteroskedasticity (GARCH) model with autoregressive moving average (ARMA) terms in the mean equation (thereby correcting for autocorrelation and heteroskedasticity) finds a positive and statistically significant correlation between the first differenced three-month U.S. T-Bill yield and the implied volatility of the Chicago Board Options Exchange (CBOE) S&P 500 Index options contract.

20The reader will note that I have been vague about whether I am talking about nominal or real rates, or policy, short-term, or long-term rates. It is possible to spin a story about each rate. For example, insurance companies with fixed nominal commitments would be concerned about nominal rates, while those with real commitments would be concerned about real rates. To the extent that policy rates can affect long-term nominal rates (by signaling the stance of future monetary policy) as well as real rates (by inducing greater risk tolerance), policy rates may indeed be relevant.

21This is not a novel observation—the failure of “portfolio insurance” strategies during the crash of 1987 was a wake-up call about the risks inherent in dynamic hedging strategies.

22Of course, hot money may well have paid up on average, by accepting lower interest rates from the emerging market when it knows it has an assured and easy exit.

23I have argued that the changed incentive structures as markets become more arm’s length contribute to increased risk taking and heighten the possibility of booms and busts. Much of the debate in macropolicy circles has, however, centered on the decline in inflation as the primary explanation. More precisely, the argument
is that quiescent goods price inflation may have created a fertile new environment for asset price booms and consequent busts. Perhaps oversimplifying the well-argued ideas of Borio (2003) Borio and Lowe (2002), and Crockett (2001, 2002, 2003) in this new environment, credit expansions are less likely to be accompanied by goods price inflation. Monetary policy that is focused on controlling short-run goods price inflation is likely to exert fewer checks on credit expansion and asset price inflation. The increase in the number of booms and busts in recent years is thus attributed, in part, to the presumed death of inflation. The two sets of explanations (incentives or the absence of goods price inflation) are not mutually exclusive, and more research is required to ascertain their relative merits.

24 A related notion is statistical provisioning implemented by the Spanish authorities. Under this, bank provisioning (set aside from capital) is anchored to an average estimated over the business cycle, creating a buffer that is run down when credit conditions worsen. This reduces the degree of procyclicality associated with capital requirements but may have milder effects on booms fed by asset price inflation.

25 Market discipline might be very useful in preventing excessive risk taking or incompetent management by specific institutions, but is less useful against booms, where market participants are themselves caught up in the frenzy.

26 The author thanks David Hsieh for providing these data.

27 Private correspondence with Antoinette Schoar.
References


