Market Discipline of U.S. Financial Firms: Recent Evidence and Research Issues

Mark J. Flannery
Flannery@ufl.edu

Stanislava Nikolova
Sniki@ufl.edu

Department of Finance, Insurance and Real Estate
Graduate School of Business Administration
University of Florida
Gainesville, FL 32611

November 24, 2003

The potential for market discipline of financial firms has attracted increasing attention from academics and supervisors. Not only have financial economists written numerous papers on the topic, but the Basel Committee on Bank Supervision identifies market discipline as one of the three “pillars” on which effective regulatory oversight should be based. After initial resistance, some supervisors have come to recognize the potential benefits of market participation in bank supervision:

> The real pre-safety-net discipline was from the market, and we need to adopt policies that promote private counterparty supervision as the first line of defense for a safe and sound banking system. (Greenspan, 2001)

Market discipline also appears to suit a financial sector in which financial institutions are increasingly complex:

> I believe that market discipline is a particularly attractive tool for encouraging safety and soundness in a rapidly evolving environment. Market discipline is inherently flexible and adaptive with respect to innovations, since market participants have incentives to change the ways that they evaluate risks as innovations are adopted. (Meyer, 1999)

In short, involving counterparties in bank supervision may reduce the costs and improve the quality of prudential oversight.

Researchers have investigated whether market information is accurate and timely, and whether it can enhance the supervisory process. In this paper, we review the recent empirical evidence on these issues. In the process, we try to summarize what is known, and to identify the most important things we still need to learn. Our charge is to concentrate on the U.S. evidence, which requires us to omit several interesting foreign studies.

**Defining market discipline**

Supervision has among its primary goals the identification and correction of problems that might lead to a financial firm’s failure. At its most basic level, therefore, supervisors seek information about their firms’ failure probabilities, which they have historically assessed via mandatory financial reporting and periodic on-site examinations. “Market discipline” describes a situation in which private sector agents – equityholders and debtholders – produce information that helps supervisors recognize problem situations and implement appropriate corrective measures. Bliss and Flannery (2001) describe two distinct components of market discipline – market monitoring and market influence. Effective monitoring occurs when investors can accurately assess changes in a firm’s condition, and promptly impound those changes into the firm’s stock and bond prices. Market influence describes the ability of market participants to affect a firm’s financial decisions. Both facets of market discipline can theoretically complement government supervision.

Berger (1991) points out that market discipline requires private investors to face costs that increase as firms undertake risks, and to take action as a result of these costs. These investor reactions fall into three broad classes. First, investors can monitor bank condition and require higher coupon rates from riskier banks. When a supervisor observes higher rates (if they are paid), she knows that at least some investors have noticed a change in the bank’s probability of failure. So monitoring creates information that supervisors can use. But what if the bank refuses to pay higher rates? This is the second class of investor reactions: at least some
investors will withdraw uninsured funds from the bank (Billett et al. (1998)). In response, the bank must raise new funds elsewhere and/or curtail its activities. This interaction generates no price signal for supervisors, but the shift from uninsured to insured sources of funds conveys information about the market’s assessment. Moreover, the withdrawal of uninsured money probably forces the institution to shrink, which is what supervisors might also prescribe for an institution whose risk has risen. So in this case investors’ actions influence the bank’s position. The third class of actions forces the bank to restore its initial default probability. Supervisors would observe neither a rate change nor a withdrawal of uninsured liabilities if the threat of either action leads the bank to restore its capital ratio – either by reducing asset size or raising (retaining) new equity. This constitutes market influence, and if managers value their debt ratings we can expect investors or the rating agencies to force at least some corrections that are consistent with what supervisors would like to have done.

Before we discuss each class of market action at greater length, we should emphasize that market discipline requires private investors to bear the costs of their bank becoming more likely to fail. Obviously, a supervisory system that provides implicit guarantees against creditor losses cannot expect to secure much benefit from market discipline.

We have relatively little empirical evidence about the ability of market participants to influence financial firms’ decisions, which is likely the result of market influence being hard to identify (Rajan (2001)). The basic problem is that market prices impound expected future actions. The impact of any shock on a firm’s value therefore depends on the shock itself and the managers’ anticipated reaction to the shock. A small decline in share value following a negative shock could mean either that the shock was small and managers were not expected to respond at all, or that the shock was large but investors anticipated that appropriate managerial action could offset part of its effect on the firm. (See also “Influencing bank capitalization?” below.)

Moreover, the decisions of a financial firm reflect both the influence of market participants and the existing regulatory framework, making it difficult to isolate the effect of either one. In an effort to draw inferences about the effectiveness of market discipline per se, researchers must compare the behavior of similar firms operating under different supervisory arrangements. For example, Nier and Bauman (2002) use a sample of banks from 32 countries and document a positive association between capital and risk after controlling for government guarantees. Gunther et al. (2000) studied Texas banks in 1910, when they had access to two alternative insurance plans. One plan charged a flat-rate premium while the other required additional security if the financial state of the bank deteriorated. The authors find that banks choosing the latter, were better capitalized and less risky. These capitalization and risk differences persisted or increased over time.

Compared with identifying effective market influence, studying market monitoring is relatively easy. Moreover, monitoring is a logical prerequisite for influence. It is therefore not surprising that most of the market discipline research has addressed three main questions about monitoring:

- Does the market accurately reflect information about financial firms’ condition?
- Does the market incorporate information in a timely manner?
- Does that information differ from the information available to supervisors?
To many financial economists, it seems odd to pose these questions so explicitly, since the efficiency of market prices has strong support for asset prices in general. Its high profile in studying financial firms may reflect an unusually high cost of mis-perceiving a financial firm’s true condition – e.g. because bank failures are unusually costly. Alternatively, banking firms may be more difficult to understand from the outside – a question evaluated by Morgan (2002), Morgan and Stiroh (2001), and Flannery et al. (2004), albeit with different conclusions. Strategically, it was also important to establish the accuracy of market monitoring in order to counteract a common supervisory impression that their inside information was generally (and necessarily!) more accurate than market prices.

Market participants’ ability to evaluate a bank’s financial condition is typically detected as an association between market prices of the bank’s claims and indicators of its asset value and risk exposures. The most common such indicators are the risk composition of the bank’s asset portfolio, its capitalization status, its regulatory or credit-agency rating, or its failure. Evidence of the market’s monitoring ability can also be found in various event studies, which document how bank claim prices respond to new information about the bank’s condition. Finally, even if the market is found to have accurate and timely information, the question still remains of whether its information can be valuable to regulators. Researchers address this question by investigating the marginal contribution of market indicators to supervisory models, or by using market indicators to forecast regulatory ratings.

**Evidence of market discipline in debt prices**

Market discipline in the debt markets can be manifested as a change in the cost of uninsured funds. As the perceived risk of a banking firm increases, holders of the firm’s uninsured liabilities will require a higher promised return to compensate for higher expected losses.

The early search for evidence of market discipline in debt prices investigated the contemporaneous relationship between bank risk indicators and subordinated debt yields or large deposit rates. A number of recent studies follow this same approach. They document that issuance and secondary-market risk premiums on traded subordinated notes and debentures are correlated with accounting measures of risk, asset portfolio composition, credit-agency or regulatory ratings, and probability of undercapitalization or of failure (Jagtiani and Lemieux (2001), Morgan and Stiroh (2001), Sironi (2002), Evanoff and Wall (2002), Jagtiani et al. (2002), and Krishnan et al. (2003)). The analysis of the rates on large uninsured CDs yields mixed results. Hall et al. (2002) focus on banks that have a satisfactory regulatory rating and document a positive relation between CD yields and measures of bank risk. On the other hand, Jagtiani and Lemieux (2001) find no evidence of market discipline in the uninsured CD market for a small sample of bank holding companies (BHCs) with failing subsidiaries. An innovative approach to the issue is offered by Furfine (2001), who examines the existence of market discipline in a debt market overlooked by previous studies – the overnight federal funds market. The evidence there is expected to be the strongest since counterparties are likely to be more sophisticated than the average investor, and loans are likely to impose significant expected losses given their size and lack of collateral. Consistent with this expectation, the author documents that interest rates paid on federal funds transactions reflect differences in credit risk across borrowers.
Another approach to investigating the existence of market monitoring is by testing whether the prices of uninsured debt claims respond to relevant news in a rational manner. If the debtholders of a financial firm are concerned with their expected losses, then they will adjust the valuation of their claims as new information about the firm’s default probability is revealed. Two recent event studies provide evidence of market discipline by documenting that the prices of uninsured financial-firm liabilities respond to news in a reasonable manner. Harvey et al. (2003) examine whether the Federal Reserve’s announcement that trust-preferred stock can be included in Tier I capital affected the value of bank debt. The authors find that following the announcement, credit spreads of banks that subsequently issued trust-preferred stock decrease. This is consistent with the hypothesis that debt markets value the default probability decrease from the addition of trust-preferred stock to capital. Allen et al. (2001) examine changes in debt valuation for BHCs that decide to convert to financial holding companies (FHC) under the Gramm-Leach-Bliley Act. They find that conversions result in lower yield spreads consistent with the lower risk likely to be attained through greater diversification.

Another recent change in the market discipline literature has been a shift from contemporaneous affirmation to forecasting. While earlier studies focused on contemporaneous associations between uninsured liability prices and indicators of firm risk, new ones explore whether information in current prices corresponds to future outcomes. This question is intuitively appealing since prices reflect expectations of future cash flows. Analysis of the forecasting ability of subordinated debenture credit spreads demonstrates that these can explain one quarter-ahead regulatory ratings better than capital ratios could (Evanoff and Wall (2001)), they can improve the forecasting accuracy of supervisory rating models four quarters prior to inspection (Krainer and Lopez (2003)), and they are more closely correlated with bank risk measures in the 12 months prior to bank failure (Jagtiani and Lemieux (2001)). Thus, prices of subordinated debentures reflect not only the current condition of financial firms, but its future condition as well.

**Evidence of market discipline in debt quantities**

Market discipline in the debt markets can be detected not only as a change in the cost of uninsured funds, but also as a change in their availability. If holders of uninsured liabilities believe that a financial firm is becoming increasingly risky, they might react by withdrawing or withholding their investment. In addition, financial firms undertaking more risk are faced with higher borrowing costs of uninsured funds and can choose to use insured funds instead. These supply-side and demand-side changes of uninsured debt quantity in response to changes in perceived risk would be evidence of market monitoring and influence respectively. Consistent with the existence of market discipline, recent studies document that as the financial condition of banks (Billet et al. (1998), Jagtiani and Lemieux (2001), and Hall et al. (2002)) or thrifts (Goldberg and Hudgins (1996), Park and Peristiani (1998), and Goldberg and Hidgins (2000)) worsens, these institutions increase their reliance on insured deposits. McDill and Maechler (2003) also ask whether uninsured depositors discipline banks by withdrawing their funds but they explicitly account for the simultaneous determination of deposit quantities and rates. When a bank’s fundamentals deteriorate, they find a shift in both its demand for uninsured
funds and the markets supply of such funds. That is, the volume of uninsured deposits at a
deteriorating bank falls, even though the bank responds by offering higher rates.

It is a wide-spread regulatory concern that the effects of market discipline on the quantity of
deposits can have undesirable consequences. One often cited reason for this concern is that
depositors will withdraw funds suddenly in a “run” to reduce looming losses, rather than doing
so gradually as changes in a bank’s risk profile occur. While some believe that “runs” can
cause even a healthy bank to fail, Calomiris and Mason (1997) find no evidence to support
such a belief. They analyze the 1932 Chicago banking panic and find that failing banks could
be distinguished from non-failing banks at least six months before the panic. They conclude
that failures reflected the relative weakness of failing banks in the face of a common asset
value shock.

Another cited reason for regulatory concern is that depositors may use information about one
bank to infer incorrectly the condition of others. One bank’s failure may thus trigger
contagious runs and wide-spread panics. Existing studies have documented that bank-specific
news events, such as announcements about the quality of a bank’s portfolio (Docking et al.
(1997)) and about formal supervisory actions at the bank (Jordan et al. (2000)), do affect the
valuation of other banks. As in Calomiris and Mason (1997), these studies find evidence of
rational, information-based spillover, as opposed to an indiscriminate contagion.

**Evidence of market discipline in equity prices**

Although most of the research has evaluated market discipline by examining debt prices or
quantities, a few studies have looked at equity prices. This limited research indicates that share
prices reflect a bank’s current condition, and can perhaps help predict future changes. Recent
event studies document that equity prices promptly impound new information about a bank’s
condition: Moody’s downgrades (Billett et al. (1998)), loan-loss reserve increases (Docking et
al. (1997)), and announcements of formal supervisory actions at the bank (Jordan et al. (2000))
result in negative abnormal equity returns. In addition, as we discuss below, equity-market indicators
can help predict changes in a firm’s financial condition one to two years before these changes
materialize (Elmer and Fissel (2001), Berger et al. (2000), Curry et al. (2001) and Krainer and
Lopez (2002)).

**Influencing bank capitalization?**

Market discipline need not manifest itself as a change in the cost or availability of uninsured
funds. Since any firm’s default probability is a function of both asset risk and leverage, a
financial firm can keep its default probability constant if it balances changes in risk with
changes in capital. The causality in the relationship between risk and capital can flow in either
direction. Banks might decide to increase risk and then build up capital to reduce the
probability of default. Alternatively, they can first accumulate capital and then choose to
substitute into riskier activities, perhaps because they need to take on more risk in order to pay
for their higher overall financing cost. An incipient increase in risk or reduction in
capitalization can generate debt price or quantity pressure to adjust bank capital to its perceived
risk exposure. Such behavior is consistent with market influence even though it can leave the quantities and prices of uninsured liabilities unaffected.

Early evidence of the interdependence between bank risk and capital can be found in Cargill (1989) and Swindle (1995). Cargill (1989) documents that CAMEL ratings are not significant in explaining interest rates on large uninsured CDs. Using essentially the same data, Swindle (1995) finds that regulatory capital ratio changes are a function of CAMEL ratings. The results in the two studies suggest that increases in risk can be compensated for by increases in equity, thus leaving the risk to debt-holders unchanged. More recently, Shrieves and Dahl (1992) and Calomiris and Wilson (1998) find that increases in bank risk are positively related to increases in bank capital. Finally, Flannery and Rangan (2003) document that the substantial capital buildup by large BHC during the 1990s is likely due to the increased risk exposure and risk aversion of bank counterparties as implicit and explicit government guarantees have decreased.

**Using market discipline to enhance bank supervision**

We find it encouraging to learn that market information is frequently mentioned in formal supervisory reports (Feldman and Schmidt (2003)). However, it is “not clearly integrated into or supportive of a supervisory assessment.” (page 1) If market assessments are to influence supervisors, we must design procedures through which market information is incorporated meaningfully into the supervisory process. This incorporation can take three general forms. First, market information may help supervisors assess a firm’s current condition. Second, it can act as an early warning signal of the firm’s future condition. Third, it can permit (or require) supervisors to address problems promptly.

Assessing a firm’s current condition should be viewed as a “forecasting” problem: given available information, what is the firm’s true default probability? Statistical theory indicates that combining two estimates will produce a more accurate assessment than either one alone, provided the estimators’ prediction errors are not perfectly correlated. So long as market information accurately reflects bank condition, on average, combining market and supervisory information should yield a more accurate assessment of the firm’s true condition.

We have already cited studies that document the relationship between market signals and bank risk. Does this market information add value to supervisory information? A number of studies show that equity-market and debt-market indicators can marginally increase the explanatory power of BOPEC or CAMEL forecasting models. Berger et al. (2000) find that supervisory assessments are less accurate than equity market indicators in reflecting the bank’s condition except when the supervisory assessment is based on recent inspections. Gunther et al. (2001) show that equity data in the form of expected default frequency can add value when included in BOPEC forecasting models. Curry et al. (2001) find that adding simple equity-market indicators (price, return, and dividend information) adds explanatory power to CAMEL rating forecasting models based on accounting information. Evanoff and Wall (2001) show that yield spreads are slightly better than capital ratios in predicting bank condition. Krainer and Lopez (2003) find that equity and debt-market indicators predict subsequent BOPEC ratings and that including these in a BOPEC off-site monitoring model helps identify additional risky firms.
Only a small subset of the above papers (Krainer and Lopez (2002, 2003) and Gropp et al. (2001)) assess the model’s forecasting ability not in terms of $R^2$ statistics alone (which treat all errors as equally important), but by examining the model’s propensity for specific kinds of errors. In the context of bank supervision, we conjecture that a Type I error – failing to identify a true problem – is more costly than a Type II error – falsely classifying a sound bank as troubled. Krainer and Lopez (2002, 2003) are encouraged when their market-augmented decision rule slightly reduces the number of false negatives, even though the market information adds little to the model’s $R^2$ statistic.

At the present time, it appears that market information adds rather little to the explanatory power of models that rely exclusively on accounting information. Given that supervisors actually know more than is contained in these accounts, this is a low threshold for judging the performance of market information. Further research in this area is clearly important. Without a substantial increase in predictive abilities, however, market information is unlikely to help supervisors identify problems at an early stage.

Even if market information cannot systematically improve supervisory assessments of current or future conditions, contemporaneous affirmation of supervisory information can still provide substantial value. Supervisory judgments can be buttressed by market data, provided that market data is properly interpreted. By re-affirming supervisors’ judgments, market information may enable supervisors to act sooner when they perceive a problem. Conversely, market information may cause appropriate forbearance if it suggests that the supervisory view is too bearish.

A second use for market information also requires only that market signals provide an unbiased contemporaneous assessment of bank condition: supervisors can be required to take specific actions when a market signal becomes adverse enough. Sufficiently accurate market signals might usefully serve as a tripwire to forestall supervisory forbearance, as investigated by Evanoff and Wall (2000). If market signals are noisy, a strict rule of this sort will generate Type I and Type II errors. Nevertheless, the cost of these errors might be lower than the cost of errors made under the alternative supervisory procedures.

**Stock or bond market indicators?**

Most of the existing literature on market discipline of financial firms has focused on information in debt prices or credit spreads, largely because the debtholders’ payoffs most closely resemble those of regulators. However, this advantage of debt market prices is balanced out by a number of disadvantages. To start with, debt prices are notoriously difficult for a researcher to collect. While some corporate bonds trade on NYSE and Amex, they account for no more than 2 percent of market volume (Nunn et al. (1986)). The accuracy of bond data is also problematic. Data on OTC trades or quotes are not collected into a single database. Indeed, Hancock and Kwast (2001) compare bond yield data from four sources and find that the correlations between different sources are only about 70-80 percent. Moreover, many bond prices are based on matrix valuation rather than on actual trades, and Warga and Welch (1993) document that there are large disparities between matrix prices and dealer quotes. Even for bond transaction prices, Saunders et al. (2002) show that the corporate bond market is
characterized by a small number of bidders, slow trade execution, and large spreads between the best and second-best price bids.

Another factor that may limit the usefulness of debt prices as a tool of market discipline is that debt issuance has been endogenous. According to Covitz et al. (2001), managers tend to issue new debentures or other uninsured claims disproportionately when market investors are relatively bullish about the firm. The result is a weaker connection between new issue rates and the bank’s true condition. This possibility has led some researchers to suggest that debenture issuance should be mandated at regular intervals (Calomiris (1998), Ferguson (1999), Kwast et al. (1999), Meyer (1999), Benink and Schmidt (2000), Evanoff and Wall (2000), and U.S. Shadow Regulatory Committee (2000)). Although there is evidence that the implementation of these proposals can be valuable, the existing literature has not assessed the cost (if any) of forcing banks to issue debt regardless of market demand for it.

Even if there were no self-selection problem and bond data were readily available and accurate, extracting risk information from debt spreads is complicated. The typical approach is to use debt prices and calculate yield spreads as the difference between a corporate yield and the yield on a Treasury security of the same maturity. After adjusting for redemption and convertibility options, sinking fund provisions, and other relevant features, spreads are assumed to measure credit risk. However, corporate yields differ from Treasury yields for a number of reasons other than credit risk (see Elton et al. (2001), Huang and Huang (2002), Delianedis and Geske (2001), and Longstaff (2004)). They include premiums for tax, liquidity and expected recovery differences between corporate and Treasury bonds, as well as compensation for common bond-market factors. This new understanding of the composition of bond spreads raises new questions about their interpretation.

Saunders (2001) points out that contingent-claim models of firm valuation imply that in perfect markets both equity and debt prices will reflect the same information about firm market value and portfolio risk. Why not extract market information from share prices? In contrast to debt markets, equity markets are liquid and deep. Accurate equity prices can be obtained frequently, and more financial firms have traded equity than traded debt. Despite these data advantages, researchers have been reluctant to embrace the idea that share prices can convey valuable information to supervisors. Junior bonds have payoffs similar to those of the federal safety net, making their price changes easier to interpret in most cases. By contrast, the value of any safety net subsidies is impounded in the bank’s share price. As a result, equity price increases do not unambiguously correspond to a lower expected claim on the federal safety net. Indeed, under some circumstances, an insured institution’s equity value can rise simply because its portfolio risk has risen – which leaves the bank’s failure probability higher than before. This calls for using equity-market indicators other than prices or returns to extract information about firm risk. Such indicators are proposed by KMV (Crosbie and Bohn (2002)) and Gropp et al. (2001), who combine equity prices, firm leverage, and historical equity volatility to construct equity-based default risk measures.

To the extent that debt (or equity) prices contain noise, or fail to conform to the Black-Scholes assumptions used to “back out” risk parameters, using both securities (where they are available) might provide more accurate information. Gropp et al. (2001) and Krainer and
Lopez (2003) document the advantage of this approach by showing that a model using both equity-market and debt-market indicators to forecast bank risk, outperforms a model using either set of indicators alone. Nikolova (2003) proposes a structural way to combine information from equity and debt prices. She observes that if a firm’s equity is modeled as a call on its assets and the same firm’s risky debt is modeled as risk-free debt short a put, then the equity-call and debt-put models can be combined in a system of equations and simultaneously solved for the market value and volatility of assets. The resulting asset volatility is entirely forward-looking. Nikolova finds that risk measures produced by this approach are more closely related to indicators of bank condition. They also have higher explanatory power in forecasting material changes in bank condition up to three quarters before these changes materialize.

Financial conglomerates: How big a cloud?

Our discussion thus far reflects an implicit assumption in the literature that the legal entity issuing a market security is the same as the entity whose safety concerns supervisors. Yet the U.S. has long maintained a distinction between “banking” and “non-banking” financial activities, and much has been written about our need to limit the activities covered by the social safety net. With the passage of the Gramm-Leach-Bliley Financial Modernization Act of 1999, this issue becomes more compelling. For a financial conglomerate combining traditional (U.S.) banking, investment banking, and insurance, which parts should the safety net protect? Market discipline will help assess the condition of protected activities only if the relevant subsidiaries issue stand-alone securities. Even then, supervisors must understand the extent to which real or imagined guarantees from the parent to the subsidiaries’ claimholders affect the stand-alone security prices. Intra-conglomerate guarantees may strip the relevant information out of the rate spread (e.g.) on subsidiary debentures. The de facto separation of subsidiaries within a bank holding company has often been discussed in the U.S. without reaching consensus. At this point in time, however, we must re-consider this issue if market discipline is to supplement government supervision of the financial system.

Summary and conclusions

Recent research into market disciplinary forces in banking has refined the idea of market discipline and how it may complement traditional supervisory methods. The initial focus on bond or deposit credit spreads has expanded to include studies of funding composition: the balance between insured and uninsured sources of funds. Researchers have also begun to recognize that equity prices may contain relevant information about a firm’s condition, even if shareholders’ interests are not always aligned with those of the safety net. We have also made progress judging the marginal value of market information in the supervisory process, moving toward more sophisticated indicators of information, such as the number of Type I and Type II errors associated with market-augmented classification rules. Assessing a bank’s condition from the outside is essentially a forecasting problem and treating it this way helps to focus the relevant questions and issues. Along with these conceptual (and empirical) advances, the increased conglomerization of financial service corporations in the U.S. raises questions about the relation between market discipline and the extent of the federal safety net.
We conclude by listing the open research and policy questions that seem most important for pushing forward the use of market discipline as a complement to more traditional methods of safety and soundness supervision.

- We need to determine the best method of extracting information from security prices and the best procedure for meaningfully incorporating this information into supervisory processes.

- We need further information about the marginal contribution of market information to the “Type I – Type II” accuracy of various forecasting models.

- Financial economists and supervisors should try to articulate an appropriate loss function for forecasting and classification models. How expensive are undiagnosed problems vs. false alarms?

- The emergence of financial conglomerates complicates supervisory uses of market discipline, and requires a re-consideration of how extensive should the federal safety net be.
References


---

1 This perverse price reaction occurs only when the bank is very highly levered. Subordinate debt shares the same feature in some situations.