Safe banking to avoid moral hazard

Received: 16th December, 2007

Sankarshan Acharya
is Associate Professor of Finance at the University of Illinois at Chicago. Sankarshan holds doctorate and masters degrees, respectively, from Northwestern University and the Indian Institute of Technology. His research includes valuation of latent assets, optimal bank regulation, CEO compensation, and arbitrage pricing of derivatives, published in leading journals like the *Journal of Finance and International Economic Review*. His policy research has been used in the Federal Deposit Insurance Corporation Improvement Act and Basel Accords of the Bank of International Settlements. He has counselled the governments of the USA, India and China.

Department of Finance, University of Illinois at Chicago, 601 S. Morgan Street, M/C 167, Chicago, IL 60606, USA
E-mail: sacharya@uic.edu

Abstract This paper argues that the moral hazard risk faced by a banking or non-banking firm can be dissipated by arbitrage trading in a market economy. This implies an optimal policy for (a) discontinuation of government insurance, regulation or bank intervention and (b) promotion of market-based safe banks that only invest in government securities and universal banks that invest in all assets. Safe banks can serve panic-prone depositors and thus minimise the systemic risk faced by an economy due to banking panics and runs. The risk premium on assets of a leveraged firm can be shown to be negatively related to asset volatility. The minimum asset-to-debt ratio threshold below which a firm goes bankrupt is an increasing function of the asset risk premium.

Keywords: *efficient resolution of moral hazard, arbitrage pricing, hedging, safe banking policy*

INTRODUCTION

The US Government guarantee of bank deposits was instituted to contain the systemic economic risk of banking panics and runs. However, it has given rise to two problems of moral hazard:

- excessive risk-taking by banks with insured deposits;
- excessive regulation to monitor insured banks.

These two problems impose significant costs on taxpayers. For example, when banks are insured but take excessive risks, the government bears the cost of bailing them out. This may impose potentially higher costs *ex post* than the deposit insurance premiums the government may collect *ex ante* from banks. The banks pass on their insurance costs to taxpayers by charging higher interest on loans or by paying lower interest rates on deposits than would be feasible otherwise. Taxpayers also bear the burden of regulatory agencies, both directly as well as indirectly as banks pass audit costs on to customers.

This paper argues that arbitrage-free hedging strategies can dissipate the first moral hazard risk. This establishes the importance of financial institutions' risk-hedging practices. The common belief that the market is supreme is thus shown as true in the sense that first-best efficient outcomes can result even when the first moral hazard may be present. This result is consistent with casual empiricism that government meddling, by and large, results in inefficient outcomes.[1]

That the first moral hazard risk can be dissipated by arbitrage trading seems fundamentally important. If the arbitrage pricing condition — not present in the extant models of moral hazard — is removed from the model analysed in Acharya (2007),[1] the first-best efficient outcomes do not obtain. This proves that arbitrage trading — which may be interpreted as the super-monitoring role of the market — can achieve the first-best efficient outcomes. The extant models on moral hazard motivate the government to intervene in the market, which is argued here to be inefficient.

Efficiently resolving the first moral hazard problem leads to alleviation of the second problem. A major implication of arbitrage trading in a real-world marketplace is the efficient discontinuation of government insurance, regulation and the intervention of banks, and the promotion of market-based safe banks and universal banks. These safe banks would invest only in government securities. Universal banks would be the same as current financial institutions, offering commercial banking, investment banking, insurance and brokerage services under a single roof. Universal banks would continue to invest in all assets.

Under the proposed market-based safe banking policy, the government would no longer insure deposits in any bank, safe or universal. The newly created safe banks would attract panic-prone depositors. The universal banks would continue to serve all types of investors, including those willing to take high risks. Regulators would need to monitor only the safe banks to ensure that they invested in the safest government securities.

Investors depositing in safe banks would thus have a total guarantee of their deposits, not through insurance, but because of the banks' investment in only the safest assets. The government would require the safe banks to maintain a minimum positive threshold capital to prevent unscrupulous managers from paying themselves excessive monies and perquisites to erode the safe investments below the amount of deposits. The safe banks would thus be quasi-governmental financial institutions chartered to contain such banking panics as erupt when deposits of panic-prone households are not insured by the government.

The proposed safe banking policy, with all deposits in safe banks fully protected, would be at least as good as the current system of partial government guarantee. More importantly, a safe banking policy would obviate the enormous risk and cost of government monitoring and regulating all banks currently offering federal deposit guarantees through the Federal Deposit Insurance Corporation (FDIC).

The current system of partial deposit insurance in all FDIC-insured banks has proven to be enormously costly to taxpayers. US taxpayers lost an estimated \$300bn to rescue the failed Savings and Loans institutions in the early 1990s. The current banking crisis may cost US taxpayers as much as \$1trn. A study by the FDIC shows that government monitoring of insured banks carries little new information. In the FDIC Improvement Act 1991, US Congress therefore mandated public ratings agencies to rate bank assets and to set bank capital and deposit insurance norms based on such ratings. Yet because public ratings agencies are paid by the bond issuers, they have faced ongoing criticism for errors in their ratings. If banks pay these agencies for the ratings of their assets, and regulatory standards are based on such ratings, then taxpayers may bear new burdens in the future.

The safe banking policy proposed here thus seems to be the best optimal course available to taxpayers. The safe banking policy stems from a very general model of a firm analysed in Acharya (2007).[1] Making that model more specific with reasonable cost functions for monitoring/auditing the assets and issuance of capital by firms can result in interesting implications:

- the asset risk premium is negatively related to the volatility of a leveraged firm;
- the asset volatility and risk premium are both increasing functions of the asset-to-debt ratio;
- the minimum asset-to-debt ratio threshold below which the firm goes bankrupt is an increasing function of the asset risk premium.

The model is on continuous time. It is dynamic. It considers a leveraged firm. Debt-holders incur costs to monitor the firm periodically. The firm incurs transaction costs to raise new equity periodically. The market acts as a super-monitor to preclude arbitrage opportunities in equilibrium. Equity-holders are assumed to maximise the market value of equity and control the choice of the standard deviation (volatility) of the return to assets of the firm.

Standard asset pricing models imply a linear relationship between the expected return and variance of all efficient portfolios. They show that the expected return to an asset is linearly related to its systematic risk in a frictionless world, where investors prefer a higher mean return, the higher the variance. But these models show no relationship between a real asset's expected return and volatility.

RESULTS AND POLICY IMPLICATIONS

Consider a representative firm with risky debt and equity over a discrete sequence of time. At any time, the firm can be closed due to either bankruptcy or liquidation. The equity-holders may raise new equity capital at any time by incurring some new capital issuance transaction cost. A negative new capital issuance can be interpreted as a special (incremental) dividend. The real-world corporate rules for bankruptcy and liquidation can be applied.

The market and the debt-holders can freely observe the amount of new capital raised. However, they have to incur a monitoring cost to observe or assess the true value of assets and the expected return on assets. The assets chosen or invented by the owners of the firm can be presumed to generate expected rates of return. The model can determine the volatility of assets endogenously by maximising the market value of equity in two cases. The cases are identified by whether or not equity-holders' choices of firm asset volatility are *ex ante* observable by the debt-holders and the market.

The extensive form of the game can be presented with indicator variables for liquidation and bankruptcy as in Figure 1.

Debt-holders face moral hazard when the risk due to the firm's choice of volatility is not observable *ex ante* as their future payoff is affected by such risk. Debt-holders can, however, incur a monitoring cost to observe the realised value of assets, whether or not they face moral hazard due to the firm's choice of asset volatility.

A statistical probability measure can be used to determine the expected value of random future payoff to an asset. The expected payoff can then be discounted by the market-required risk-adjusted cost of capital (equal to riskless rate plus risk premium) to obtain the current fair market value of the asset. The literature on financial intermediation shows that, if lenders monitor a borrowing firm only when the firm defaults and if the cost of such *ex post* monitoring is positive, the firm will voluntarily disclose its true payoffs whenever it defaults.[2–4] Thus, the other 'moral hazard' due to the possibility of disclosing information different from the true value of assets (payoffs) on each date is resolved within the model due to costly *ex post* monitoring. The debt-holders can monitor the firm, even when it does not default, because of the moral hazard due to the asset volatility choice and to provide the market with credible information about asset value. If debt-holders monitored only when the firm defaulted, firms would not necessarily disclose their payoffs when they did not default, which would not solve the moral hazard problem.

This standard valuation procedure is not feasible for derivatives because the risk is not stationary and therefore not measurable. Derivative assets are, therefore, priced by arbitrage.

To price assets by arbitrage, a risk-neutral pricing measure is constructed, equivalently to the statistical probability measure. The risk-neutral pricing measure can then be used to determine a weighted value of the random future payoffs to the asset. The weighted value is known as the certainty equivalent of the expected payoff to the asset. The certainty equivalent so obtained can be simply discounted by the riskless rate to

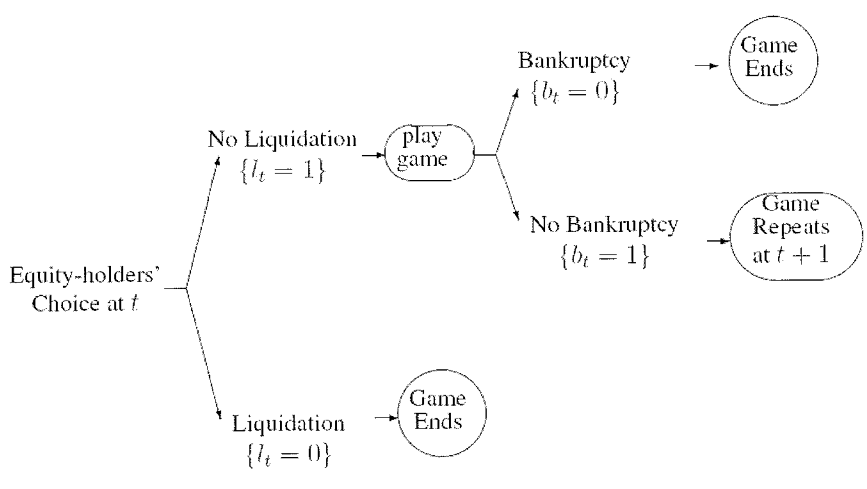

Figure 1: Equity-holder-debt-holder game

find the current fair market value of the asset.[5] If assets trade at their fair market values so determined, no arbitrage opportunities can exist, ie no non-positive current investment in an asset can yield a positive future value.

A frictionless setting within the model will obtain when the costs of monitoring and raising capital are zero. Then the equivalent pricing measure will be determined in the standard way.[8,9] In this case, the instantaneous risk premium on the asset under the original statistical probability measure is equal to the drift of the discounted asset values. The transformation under the equivalent pricing measure results in certainty equivalents of these discounted asset values with no drift.

The frictional costs for a firm which has not been bankrupt or liquidated at any time include the following:

- the equity-holders raise new capital by incurring some cost;
- the debt-holders are paid a periodic coupon amount to lend the firm;
- the firm's current investment of net assets can be observed freely, but the debt-holders can observe the payoff resulting from the investment only by incurring a monitoring cost;
- the firm may be liquidated at the end of a period;
- the market uses an appropriate measure to price assets.

First result

The risk premium on investment in debt and equity of the firm, operating in a real world with frictional costs, is less than the risk premium feasible in a frictionless (costless) world.

The leverage, volatility and drift choices are irrelevant to the market-value maximising equity-holders in a frictionless market, consistent with the Modigliani-Miller theorem.[8] A market with friction gives rise to several critical and interesting problems. Even if the volatility risk is observable at no cost (moral hazard is absent), assets cannot be traded when the asset value and risk premium are unobservable. If, however, the firm's debt and equity values can be determined in some equilibrium, an indirect trading on assets is feasible. But this involves transaction and monitoring costs of observing, whether or not there is moral hazard. The risk premium must be so determined that the market could continue to price assets with no arbitrage.

The first result motivates an important point that costs associated with frictions like monitoring and capital issuance can effectively reduce the return or risk premium. This risk premium can be determined when (i) moral hazard defined by a situation in which the asset volatility of a firm is not observable *ex ante*; and (ii) there is no moral hazard when the asset volatility can be observed *ex ante*. The main point made by this result is that the market determines an appropriate pricing measure by accounting for the monitoring and capital issuance costs as well as for the rational behaviour of equity-holders and debt-holders.

The first result shows a reduction in the risk premium due to the costs associated with friction. A lower risk premium implies a lower market perceived risk. This result thus reflects the real world where the information generated from monitoring and capital issuance tends to refine the uncertainty and thus reduce the risk perceived to be associated with the firm.

This result indicates that the market sufficiently raises the current value of the future payoff to an asset by framing the pricing measure appropriately in order to compensate for the costs associated with friction. It shows that the current market value of an investment in the assets of the firm facing friction or no friction is the same. Only the risk premium is adjusted through a refinement of the market-perceived risk due to monitoring and capital issuance. The adjustment is achieved via an appropriate pricing measure and the associated state prices that account for the costs of friction.

The first result indicates that higher monitoring and new capital issuance costs imply lower expected returns on the assets of a firm. This can be practically used for trading in the real world. Form two ratios for each publicly traded company using quarterly time series data: monitoring (auditing) cost divided by the market value of assets and new capital issuance cost divided by the amount of capital raised. Take the difference between either firm-specific ratio and the average of the corresponding ratio over all firms in the dataset. These differences represent the excess firm-specific costs. Use these differences as two independent factors to estimate the risk weights (slope coefficients) in a cross-sectional time series regression of the market model with the quarterly returns on firms' assets as the dependent variable. Proposition 1 implies that these risk weights will be negative. If the actual market price of assets deviates significantly from that implied by these risk weights, then one can trade in the equity and debt of the firm to generate profits.

Clearly, in a world with friction, monitoring plays an important role in revealing the true state of the firm, defined by the current asset value and expected return. If these states can be observed at no cost, the total market value of an investment in the firm could be incrementally higher than otherwise. The net investment in assets is carried forward.

The capital market plays the role of a super-monitor in the principal–agent problem, with the debt-holders as the principal and the equity-holders as the agent. The debt-holders enforce bond covenants via their bankruptcy and coupon requirement strategy. Conditional on the firm remaining in operation, the principal (debt-holders) 'instructs' the agent (equity-holders) to make choices on capital issuance and asset volatility in order to be compatible with the agent's (equity-holders') incentive of maximisation of the market value of equity. An equilibrium without this incentive compatibility requirement, which can be enforced when asset volatility can be freely observed, is called the 'first-best'. The principal–agent literature (eg see Harris and Raviv, 1979[9] and Holmstrom, 1979[10]) addresses the issue of obtaining a second-best solution by using the incentive compatibility constraint, ie when the agent's chosen control variables are not *ex ante* unobservable. To focus on the market's super-monitor role, the equity-holder-manager conflict is abstracted away. This also abstracts away from the potential role of the managerial labour market in resolving such conflicts.

The most appropriate interpretation of such a model is that the market acts as a super-monitor to enforce preclusion of arbitrage profits via no arbitrage conditions for equity. These conditions are not used in a standard principal–agent problem. These conditions fundamentally distinguish this model from the extant models of agency theory.

In an arbitrage-free equilibrium, the market value of equity is equal to the true value of equity, evaluated via the mutually optimal equity-holder and debt-holder strategies, and from the market value of debt and asset risk premium.

The rules for sharing between the principal and agent in the model, as determined by the super-monitor, are the risk premium on equity of equity-holders (agent) and the risk premium on debt of debt-holders (principal). Equity-holders and debt-holders will then receive compensations in the form of risk premiums, which are consistent with their systematic risk of investment in the firm.

The systematic risk-expected rate of return relationships are a consequence of the no arbitrage pricing[2] relationships. Observe, however, that the idea for a competitive managerial labour market can perhaps be viewed as an arbitrage-free pricing of the managerial marginal product of labour. It seems that this idea can be modelled to show that the first-best and the second-best will coincide even in a problem where the principal (equity-holders) does not observe the manager's efforts, which affects the payoff of both the principal and the manager. This will involve establishing an arbitrage-free market for the managerial marginal product of labour. These sharing rules are consistent with the expected instantaneous asset return, which is also determined in equilibrium. They may appear unconventional, but they are strikingly similar to standard managerial wage contracts. For example, by a linear wage contract, a manager's expected wage is equal to a sum of his wage from an alternative employment, the cost of implementing efforts and a risk-premium.[11,12]

To show that the first-best and the second-best coincide in an arbitrage-free economy, Acharya (2007)[1] considers the following problem under no moral hazard: where debt-holders' and equity-holders' objectives are simultaneously maximised, subject to the no arbitrage conditions. The bankruptcy $\{b_t = 0\}$ and the liquidation $\{l_t = 0\}$ strategies are then determined in equilibrium.

In either the problem with moral hazard or without moral hazard, the debt-holders' decision on bankruptcy and the equity-holders' choice for liquidation can be obtained in equilibrium defined by capital issuance, volatility, debt, coupon and risk premium.

Second result

The equilibrium solution for the problem with moral hazard is the same as that for the problem without moral hazard. This means first-best efficiency obtains due to arbitrage trading. In equilibrium, trading in debt generates no arbitrage opportunities, but yields an expected rate of return that is consistent with the risk of debt. In the language of a standard principal–agent model, the principal (debt-holders) receives a fixed payoff equal to the market value of debt. If this cannot be guaranteed, debt-holders will appeal for bankruptcy. Trading in debt at equilibrium market prices generates no arbitrage opportunities, but yields an expected rate of return that is consistent with the risk of debt.

The debt-holders can generate higher than the liquidation value of debt only if the no arbitrage condition is eliminated. In that case the first-best (efficient) solution will not obtain in equilibrium. If the principal freely observes the asset volatility, the results can be very disturbing, similar to Mirrlees (1974).[13] This is because the principal will set as high a coupon rate as possible and ask the equity-holders to choose as low an asset volatility as possible, making the equity worthless and the firm non-viable. All these unrealistic possibilities are absent when the market determines prices of securities (asset, debt and equity) by precluding arbitrage profit opportunities in equilibrium.

In an arbitrage-free equilibrium, the monitoring cost and capital issuance costs are ultimately borne by the firm. The equity-holders' choices therefore increase the market value of equity incrementally to compensate equally for the expected discounted costs of monitoring and capital issuance. This means that the market acting as a super-monitor compensates the debt-holders and equity-holders with an incremental market value equal to the minimum possible discounted expected monitoring and equity issuance costs. As proved in Acharya (2007),[1] this incremental market value obtains via the equivalent state pricing measure.

Third result

The third result defines a frontier (graph) of the risk premium versus volatility for every asset-to-debt ratio greater than a minimum threshold over which the firm remains open. The frontier relates the instantaneous expected asset return to the volatility of the firm. The firm is indifferent among all pairs on this frontier. Acharya (1990)[18] analyses a similar problem for banks with a fixed risk premium, resulting in a unique solution, which includes a unique volatility and a unique asset-to-debt ratio threshold below which a bank is optimally closed.

There exists a unique value of the new capital issuance rule that maximises the value of equity conditional on continuance of the firm and on asset value, volatility and risk premium. The continuing firm's equity value is a monotonically decreasing function of asset volatility given the asset value, new capital and risk premium. The equity value is a monotonically increasing function of the risk premium, given asset value, new capital and volatility. The net value of equity — the value from continuance minus that from liquidation — is monotonically (i) decreasing in volatility, given asset value and risk premium, (ii) decreasing in risk premium, given asset value and volatility, and increasing in asset value, given volatility and risk premium. The minimum threshold asset-to-debt ratio above which a firm remains open is a function of the risk premium. If the firm remains in operation, there exists an equilibrium volatility function that generates a zero incremental equity value from continuance as compared with the liquidation value. The risk premium and the corresponding volatility are functions of the current asset value.

The equilibrium expected return–volatility frontier seems very interesting, because it describes a cross-section of firms (including banks) whose expected return on assets depends on the volatility they choose. This result appears to be an economically significant departure from the standard paradigm in which the expected return to assets and volatility are presumed to be unrelated. The model with no friction and no moral hazard reduces to the standard paradigm that yields the capital asset pricing model or the arbitrage pricing theory as cross-sectional relationships between expected returns and systematic risks (betas). This frictionless model does not, however, imply any relationship between expected return and volatility. Indeed, the asset volatility is indeterminate if there is moral hazard and no friction as in Chan, Greenbaum and Thakor (1991).[15]

A very interesting endogenous relationship between the expected return and the volatility choice of a firm obtains in an arbitrage-free market with friction and moral hazard in the game between debt-holders and equity-holders.

The practical implications of the third result are profound. The algorithm for the moral hazard problem — faced by debt-holders due to equity-holders' control of the *ex ante* unobservable asset volatility — is the same as that for the absence of moral hazard. This robust result obtains because the model incorporates many real-world factors like arbitrage pricing and trading, monitoring by debt-holders, new capital issuance by equity-holders, and gaming between equity-holders and debt-holders with the market behaving like a super-monitor. Negative new capital is dividend payout. The existing models on moral hazard do not incorporate such real-world factors and so obtain inefficiency or second-best outcomes due to moral hazard. If arbitrage trading is not allowed within the model in Acharya (2007),[1] one would obtain inefficient outcomes.

The third result demonstrates that the moral hazard problem is efficiently resolved due to arbitrage trading and pricing in the marketplace. The existing belief among economists that inefficiency is natural due to moral hazard has induced the government to intervene, for example, in the banking sector. Government intervention imposes inefficiency. The current presumption is that the inefficiency of government intervention is less costly than moral hazard.

The third result makes such a presumption unwarranted by establishing that the market can act as a super-monitor to achieve efficiency. This result thus obviates calls for government intervention to supplant the efficient act of the market.

Most existing research in economics, except on the moral hazard, assumes that the market is supreme. The third result, however, shows that the market is indeed supreme in the sense that arbitrage trading leads to efficiency and first-best outcomes that cannot be achieved otherwise. This result thus shows as true the common belief in the developed world that the market should be supreme to achieve efficient outcomes. Inefficient outcomes in models of moral hazard have been the basis for government intervention because, as the current argument goes, such intervention imposes a lower tax on society than the cost of inefficiency due to moral hazard. The third result obviates such arguments.

SAFE BANKING POLICY

A major direct implication of the third result is that government insurance, regulation or intervention of banks are suboptimal due to inefficiency. An indirect implication not modelled in this paper is that sovereign countries should allow their currencies to trade freely to achieve first-best (efficient) outcomes for the global society. Smaller developing economies may, however, find it optimal to determine exchange rates by balancing the costs of banking and social instability with the benefits of job creation.[16]

The US Government now insures bank deposits to circumvent systemic risks due to panic-driven bank runs. However, the government deposit insurance can lead to moral hazard, which is defined as banks using insured deposits for excessively risky bets.

Consider, for example, a bank starting with $8 of common equity and $92 of insured deposits, which then creates a firewalled subsidiary with equity of $100 infused from the parent bank. The subsidiary then borrows $900 of insured deposits. The parent bank thus meets the statutory minimum required capital-to-assets ratio of 8 per cent, while its subsidiary has 10 per cent. On a consolidated basis, however, the bank has only $8 of equity for $1,000 of assets — or 0.8 per cent capital-to-assets ratio — and is thus badly under-capitalised. The executives of this bank may be tempted to bet the insured funds in the firewalled subsidiary on very risky investments. Such risky bets can generate high returns to the parent bank's shareholders and executives, although with a small probability, but subject the insured deposits to the risk of losses to be borne by taxpayers.

The current banking crisis is primarily confined to under-capitalised banks with risky firewalled subsidiaries, known as conduits, trusts or super-investment vehicles. The government is currently obligated to monitor and regulate insured banks to alleviate moral hazard.

The current supposition is that bank deposit insurance and regulation are less costly to society than the systemic burden due to panic-driven runs. Such supposition is, however, redundant when the market can achieve first-best efficiency through trading, as shown by the third result.

The banking moral hazard stems from excessive risk-taking, infrequent trading of assets within firewalled subsidiaries and illiquidity due to the government's restrictions on auctioning off regulated bank assets. The large insured US banks have perhaps accumulated too much risk in their firewalled subsidiaries due to a tacit government policy that such banks are too big to fail. In the current predicament, the regulators are reluctant to let banks auction off their risky assets at 'throw away' prices because of potential reprisal from taxpayers.

Efficient deregulation induced by arbitrage trading leads to a safe banking policy by which the government (i) stops insuring bank deposits and (ii) induces the market to open 'safe banks', that invest only in government securities to cater to the demand of panic-prone depositors, and 'universal banks' for others willing to take risk. Under the safe banking policy, bank deposits would not be insured by the government and banks would not face regulation or intervention beyond the monitoring of 'safe banks' to ensure their investments were only in government securities that maintained some minimum threshold capital to avoid unscrupulous management withdrawal of funds through excess payments and perquisites. The safe banking policy thus obviates capital requirements due to a bank's internal loss estimate, as argued in Acharya (1990)[17] and subsequently adopted in the Basel II agreement.

The safe banking policy would thwart banking panics as well as the current policy of insuring a limited amount of each individual's deposits in insured banks. More importantly, this policy would allow the market to beget efficiency in banking without the cost of regulation and intervention. Such a system of safe banking would circumvent systemic risks due to banking panics and runs (as well as the current system of government insurance) and allow markets to effect efficiency in banking without regulation. Safe banking would thus be optimal for taxpayers, as argued in Acharya (2003).[18] (The top management of a major US bank once was incensed when the author asked for the maintenance of the regulatory capital requirement on a consolidated basis.)

Trading has practically helped the market in the efficient collapse of the largest universal bank (Drexel Burnham Lambert) in the late 1980s, liquidation of the largest hedge fund in the late 1990s, dismantling of several large hedge funds in 2007 and recapitalisation of major banks in 2007. The safe banking policy thus seems quite robust to serious financial shocks with respect to bank failures in the economy.

Safe banks — that invest only in government securities — obviously cannot eliminate runs of the type recently witnessed on insured banks. Safe banks can only be as effective as the government-insured banks in thwarting runs. Nonetheless, the proposed safe banking policy is more efficient than the current policy of government insurance, regulation and intervention.

Fourth result

Given the asset value per dollar of debt, asset volatility is a monotonically decreasing function of asset risk premium. Given the risk premium, asset volatility is a monotonically increasing function of asset value. There exists an equilibrium risk premium as a function of asset volatility and value. The risk premium is a monotonically decreasing function of asset volatility, given asset value. The risk premium is a monotonically increasing function of asset value, given asset volatility.

The fourth result presents an interesting testable inverse relationship between asset return and volatility. The model does not explicitly establish relationships between volatility and return of portfolios or common stocks. Yet, the asset volatility-return relationship can be exploited to derive a similar negative relationship between common stock volatility and return by making adjustments for volatility of debt.

The fourth result also shows that, as firms become more stable with the stability indicated by the asset-to-debt ratio, they are able to take greater asset volatility risk. The intuition for this result is that more stable firms expect to incur lower monitoring costs, given volatility. The expected return would increase if the expected monitoring cost decreased due to a rise in the asset-to-debt ratio, unless the volatility increased sufficiently to increase the expected monitoring cost and keep the expected return unchanged. This means that, given the expected return, volatility increases with the asset-to-debt ratio of a firm. Further, the expected return to assets will increase with the asset-to-debt ratio of a firm, given volatility. These implications also offer very interesting and novel tests of the theory. Casual empiricism shows that more stable firms experience greater momentum in their prices. It is perhaps because they can afford to take greater risks, as implied by the fourth result.

The third result shows that the firm will be closed in equilibrium if its asset-to-debt ratio falls below a minimum threshold, which is a function of the asset risk premium. The fourth result indicates that this threshold is an increasing function of the asset risk premium. This is because the market uses the risk premium to discount the future payoff to the asset. A higher risk premium lowers the current value of the future payoff. This prompts the debt-holders to set a higher minimum asset-to-debt threshold for higher risk premiums on assets. This result contrasts the fixed thresholds of the asset-to-debt ratio for bank closure derived in Acharya and Dreyfus (1989)[19] and Acharya (1990),[14] where the asset volatility is assumed to be given (cf. Diamond, 1984[3] and Smith and Warner, 1979[20]).

The threshold asset-to-debt ratio here defines the equilibrium bond covenants that debt-holders would impose on equity-holders in an arbitrage-free equilibrium. As a practical matter, the expected rate of return on assets can be estimated from data as the cost of capital of assets. The threshold asset-to-debt ratio that debt-holders can design for seeking a firm's bankruptcy can be set as a decreasing function of this estimated cost of capital. This seems to have a very useful implication for bank regulatory policies within the current government deposit insurance regime. Suppose that a bank regulator chooses to act as a surrogate for bank debt-holders. Black, Miller and Posner (1978)[21] lucidly discuss the similarity between the lender–borrower relation and the regulator–bank relation. Although they do not analyse a model with moral hazard, they argue that both the borrower and the lender have an interest in minimising 'administrative costs of lending'. Administrative costs within the present model would be the expected monitoring and equity issuance costs. Acharya (2003)[14] argues that safe banking would obviate bank regulation in the best interest of taxpayers.

Here the regulator would mimic the model's equilibrium debt covenants obtained for a deregulated industry operating in an arbitrage-free market economy. The covenants are the (i) minimum asset-to-debt ratio threshold below which to close banks and (ii) debt/deposit insurance premium. The regulator would insure bank debt and deposits for a fair price equal to the equilibrium coupon on debt in the model. The minimum asset-to-debt ratio threshold below which banks should be closed would be an increasing function of the bank asset risk premium, as implied by the fourth result. The equilibrium bank regulatory policy, which would prevail in an arbitrage-free economy with deregulated banks in a deposit insurance regime, is the same as the one that the surrogate regulator would find optimal in the context of the model.

By imposing the cross-sectional capital asset pricing model on the results, the risk premium is an increasing function of systematic risk. Thus a higher bank systematic risk would imply a greater minimum asset-to-debt ratio threshold below which the bank should be closed, if the regulatory policy is to mimic an unregulated banking industry in an arbitrage-free economy.

The result that the minimum asset-to-debt ratio threshold is monotonically increasing with the volatility risk of a firm offers another novel test of the theory. Everything else being equal, operating firms with higher volatility risk have a higher asset-to-debt ratio (lower leverage) than those with a lower asset-to-debt ratio in an arbitrage-free economy. One can design trading strategies based on deviation of market prices of assets from their fair values implied by the frontier of asset values and volatilities constructed for a cross-section of all publicly-traded firms.

CONCLUSION

The market is an efficient super-monitor, notwithstanding moral hazard. Arbitrage trading induces efficient discontinuation of government insurance or regulation or intervention of banks and promotion of market-based safe banks, which invest only in government securities, and universal banks that invest in all assets, with the safe banks catering to panic-prone depositors and the universal banks to risky ones.

Acknowledgment

The author is especially thankful to anonymous referees and editors. The author is, however, responsible for any remaining errors.

References

1 Acharya, S. (2007) 'Efficient resolution of moral hazard due to arbitrage trading: Risk premium, volatility and leverage', available at http://www.pro-prosperity.com/Research/moralhazard.pdf (accessed 13th December, 2007).

2 Williamson, S. D. (1986) 'Costly monitoring, financial intermediation and equilibrium credit rationing', *Journal of Monetary Economics*, Vol. 18, pp. 159–179.

3 Diamond, D. (1984) 'Financial intermediation and delegated monitoring', *Review of Economic Studies*, Vol. 51, pp. 393–414.

4 Ramakrishnan, R. T. S. and Thakor, A. V. (1984) 'Information reliability and a theory of financial intermediation', *Review of Economic Studies*, Vol. 51, pp. 415–432.

5 Acharya, S. and Madan, D. (1997) 'Can a riskless asset be presumed to exist? A latent interest rate theory and evidence', *Chicago Board of Trade Research Symposium Proceedings*, Fall, Chicago Board of Trade, Chicago.

6 Harrison, J. M. and Kreps, D. M. (1979) 'Martingales and arbitrage in multiperiod securities markets', *Journal of Economic Theory*, Vol. 20, pp. 381–408.

7 Harrison, J. M. and Pliska, S. (1983) 'A stochastic calculus model of continuous trading: Complete markets', *Stochastic Processes and their Applications*, Vol. 15, pp. 313–316.

8 Modigliani, F. and Miller, M. H. (1958) 'The cost of capital, corporation finance and the theory of investment', *American Economic Review*, Vol. 48, pp. 261–297.

9 Harris, M. and Raviv, A. (1979) 'Optimal incentive contracts with imperfect information', *Journal of Economic Theory*, Vol. 20, pp. 231–259.

10 Holmstrom, B. (1979) 'Moral hazard and observability', *Bell Journal of Economics*, Vol. 10, pp. 74–91.

11 Holmstrom, B. and Milgrom, P. (1987) 'Aggregation and linearity in the provision of inter-temporal incentives', *Econometrica*, Vol. 55, pp. 303–328.

12 Acharya, S. (1992) 'Maximization of the market value of a firm to choose policies for hiring, compensation, firing and tenuring of managers', *International*

Economic Review, Vol. 33, No. 2, pp. 373–397.

13 Mirrlees, J. (1974) 'Notes on welfare economics, information and uncertainty', in McFadden, D. and Wu, S.-Y. (eds) 'Essays on Economic Behavior Under Uncertainty', North-Holland Publishing Co, Amsterdam.

14 Acharya, S. (1990) 'Bank capital, risk, closure, and deposit insurance premium in a game between banks and regulators', Working paper, New York University, June.

15 Chan, Y.-S., Greenbaum, S. and Thakor, A. V. (1991) 'Is fairly priced deposit insurance possible?', *Journal of Finance*, Vol. 47, No. 1, pp. 227–246.

16 Acharya, S. (2008) 'Optimal exchange rate beyond purchase power parity', *Journal of American Academy of Business*, Vol. 12, No. 2.

17 Acharya, S. (1990) 'A simple loan loss forecast rule', Memorandum to Board of Governors of the Federal Reserve System, available at: http://www.pro-prosperity.com/Research/BanksInternalIncentiveCompatibleLoanLossEstimate.pdf (accessed 13th December, 2007).

18 Acharya, S. (2003) 'Safe banking', *Journal of American Academy of Business*, Vol. 3, No. 1–2.

19 Acharya, S. and Dreyfus, J. F. (1989) 'Optimal bank reorganization policies and pricing of federal deposit insurance', *Journal of Finance*, Vol. 44, No. 5, pp. 1313–1333.

20 Smith, C. W. and Warner, J. B. (1979) 'On financial contracting: An analysis of bond covenants', *Journal of Financial Economics*, Vol. 7, pp. 117–161.

21 Black, F., Miller, M. H. and Posner, R. A. (1978) 'An approach to the regulation of bank holding companies', *Journal of Business*, Vol. 51, pp. 379–412.