The Recent Crisis: Lessons for Islamic Finance

Abbas Mirakhor and Noureddine Krichene

2009

Online at http://mpra.ub.uni-muenchen.de/56022/
MPRA Paper No. 56022, posted 20. May 2014 18:33 UTC
The Recent Crisis: Lessons for Islamic Finance

Lecture by Dr. Abbas Mirakhor* | Former Dean
International Monetary Fund’s Executive Board

*This paper is co-authored with Noureddine Krichene.
2009  Kuala Lumpur

IFSB 2nd PUBLIC LECTURE ON
FINANCIAL POLICY AND STABILITY

The Recent Crisis:
Lessons for Islamic Finance

Lecture by Dr. Abbas Mirakhor* | Former Dean
International Monetary Fund’s Executive Board

* This paper is co-authored with Noureddine Krichene.

ISLAMIC FINANCIAL SERVICES BOARD
ABOUT THE ISLAMIC FINANCIAL SERVICES BOARD (IFSB)

The IFSB is an international standard-setting organisation that promotes and enhances the soundness and stability of the Islamic financial services industry by issuing global prudential standards and guiding principles for the industry, broadly defined to include banking, capital markets and insurance sectors. The standards prepared by the IFSB follow a lengthy due process as outlined in its Guidelines and Procedures for the Preparation of Standards / Guidelines, which involves, among others, the issuance of exposure drafts, holding of workshops and, where necessary, public hearings. The IFSB also conducts research and coordinates initiatives on industry-related issues, as well as organises roundtables, seminars and conferences for regulators and industry stakeholders. Towards this end, the IFSB works closely with relevant international, regional and national organisations, research / educational institutions and market players.

For more information about the IFSB, please visit www.ifsb.org
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II.</td>
<td>The Functions and Operations of the Financial System</td>
<td>4</td>
</tr>
<tr>
<td>III.</td>
<td>The Crisis, and Explanations of its Causes</td>
<td>23</td>
</tr>
<tr>
<td>IV.</td>
<td>The Islamic Financial System and Lessons of the Recent Crisis</td>
<td>60</td>
</tr>
<tr>
<td>V.</td>
<td>Conclusions</td>
<td>68</td>
</tr>
</tbody>
</table>

REFERENCES 72
The Recent Crisis: Lessons for Islamic Finance

Abbas Mirakhor and Noureddine Krichene

I. Introduction

The current view defines financial crisis as the collapse of the financial asset market. It is said to be triggered when asset markets no longer validate expectations of continuously rising returns on financial assets. Invariably, financial crises have involved the banking systems. Signs of the crisis are said to be: (a) large economic imbalances, particularly significant current account deficits; (b) asset price inflation; (c) rising leverage; and (d) slowing pace of economic growth. The consequences of financial crisis are considered to be: (a) a large reduction in income and employment; (b) a substantial rise in the growth of public debt; and (c) a prolonged period of recovery (Eichengreen and Mitchener, (2003); Reinhart and Rogoff, 2009). There is now a consensus that the current global financial crisis was triggered by the collapse of asset prices in the sub-prime mortgage sector of the US financial sector, which then became the epicentre of the global crisis. For this reason, analysis of the crisis has to start with the financial structure of the latter system and the conditions that led to the crisis.

Two explanations of the crisis have emerged. A conventional explanation and an alternative one. The first sees the emergence of the crisis as a consequence of large global macroeconomic imbalances that led to large savings by the emerging markets. This was itself a

* The authors gratefully acknowledge the valuable assistance of their colleague Ms. Aida Hidayah.
consequence of the financial crises of 1997–2000 in the emerging markets that prompted these countries to emphasise exports and motivated the accumulation of large primary surpluses. The resulting savings were accumulated as large central bank reserves, which were then invested in government bonds issued by industrial country governments, especially the United States. In turn, this led to low medium- to long-term interest rates, a large expansion of credit and debt, and rapid expansion of substantially large liquidity in a ferocious search for yield. The latter led to an aggressive incentive structure for the promotion of financial innovations and engineering of complex instruments with risk–return characteristics that appeared more attractive than the risk exposure attributes of the underlying assets. This process encompassed the whole spectrum of activity, design, origination, packaging, trading, distribution, wholesale and retail. Increased global demand for financialised assets led to higher prices for these instruments, thus validating expectations of ever-increasing asset prices. Then came the collapse of asset prices in the real estate market in the United States which, due to the interconnectedness of the asset markets worldwide, spread rapidly and globally through the contagious process.

The alternative view of the cause of the crisis holds that instability and crises are inherent and unavoidable in financial capitalism, because the structure of firms and households is inherently fragile, and the emergence and confluence of certain conditions convert fragility into instability and crises.

The Islamic financial system represents a financial structure substantially different from the dominant paradigm. It is supported by the network of institutional elements whose existence creates a strong barrier to the emergence of instability. This system is only in its nascent stage of development, however, and is operating in coexistence with the conventional system in a hybrid form in which many of its supportive
institutional elements either do not exist or are weak and incomplete. While theory has demonstrated that an Islamic financial system would be more stable than the dominant system, in practice such stability cannot be taken for granted without the full operation of its supportive institutional framework. This paper contends that the emergence and consequences of the recent financial crisis in the dominant system hold valuable lessons for the operation of the existing Islamic finance. Specifically, the paper suggests that because, in its present form, Islamic finance is embedded in systems that share many of the institutional features of the dominant system, to ensure its survival – until the emergence and operationalisation of its own supportive institutional framework – Islamic finance requires, as a matter of urgency, a regulatory-supervisory framework much stronger than exists in the dominant system. In particular, such a framework has to be unified, uniform and multinational, covering all the economies in which, in one form or another, Islamic finance has been adopted. Such a framework has to be comprehensive and legislatively based, and include all facets of the financial system. A regulatory-supervisory system of this kind is far more important to the Islamic financial system in its present stage of evolution than is the case with the dominant system, because the failure of even a few Islamic finance institutions would pose a far greater systemic risk and cause more reputational damage than would similar failure in the dominant system.

To render this paper self-contained, Section II presents a brief definitional survey of the dominant financial system, its structure and functions, and its important constitutive elements. Section III provides a summary description of the current crisis, as well as the conventional and alternative explanations of its causes. Section IV discusses the elemental features of an Islamic financial system that distinguish it from the dominant system, as well as its stability characteristics. This section also presents the major elements of the institutional framework that support the Islamic financial
system and its stability. Finally, Section V concludes the paper by making a case for the design and implementation of a universal, comprehensive, legislatively-based regulatory and supervisory framework that ensures minimisation of systemic risk to the Islamic financial system at this stage of its development.

II. The Functions and Operations of the Financial System

An essential feature of market economies is the existence of intermediaries: firms that facilitate transactions between producers and their suppliers of input, on the one hand, and between producers and consumers, on the other. Spulber (1996, p. 135) defines an intermediary as “an economic agent that purchases from suppliers for resale to buyers, or that helps buyers and sellers meet and transact. Intermediaries seek out suppliers, find and encourage buyers, select buy and sell prices, define the terms of transactions, manage the payments and record keeping for transactions, and hold inventories to provide liquidity or availability of goods and services.” A financial system can be thought of as a collection of financial firms and institutions that deal in and with functions of payments clearing and settling; financial resource pooling and allocating; transferring and transforming financial resources across time, space and kinds of assets; managing risk; and collecting, analysing and providing information, as well as dealing with informational problems. A modern financial system is comprised of banks, non-bank financial institutions, firms in securities markets for money, debt and equity instruments, insurance companies, credit bureaus, credit rating agencies, managed funds, institutional investors, dealers and brokers, and a central bank. These constituent elements perform a wide range of services which, inter alia, include: creating and providing liquidity; reducing the costs of financial transactions; screening and monitoring the creditworthiness and payment performance of their clients; bringing together buyers
and sellers, lenders and borrowers; providing insurance against risk; and managing the economy’s payments system.

In short, a financial system is a collection of institutions that facilitate transactions between financial resources and their users, and provide support for the real sector of the economy to convert primary resources into production for final use (Fry, 1995). In other words, the institutions comprising a financial system intermediate between surplus finance units and the real sector investors. Because uncertainty regarding the future is a fact of life, there are risks intrinsic to the act of investing on the part of those providing financing. It is believed that, generally, surplus fund holders wish to avoid risk, while entrepreneurs are by nature risk-takers. Investment projects can fail due to no one’s fault. But they can also fail because of fraud, misrepresentation, negligence, and the provision of incomplete or wrong information. These failures arise due to an informational problem referred to as “asymmetric information”. Asymmetric information exists when any one party to a transaction has information regarding the subject matter of the transaction that the other party does not possess. It can appear as adverse selection or moral hazard. The first appears when the wrong transaction is selected because of hidden or inadequate information. Moral hazard is a concept that originated in the insurance field and refers to the tendency of the insured to make less effort to avoid risks than they would if they had no insurance. In financial transactions, the concept refers to a situation where the entrepreneur seeking financing intends to use the funds differently than agreed upon with the surplus funds holder, who either has insufficient information regarding the entrepreneur’s intention or has no control over the entrepreneur’s behaviour to mitigate the risk of moral hazard. This leads to the need for monitoring mechanisms that ensure the entrepreneur’s behaviour is compatible with that expected at the time of the transaction. The associated costs are referred to as “monitoring costs” (Fry, 1995, pp. 305–312).
It is thought that because of monitoring costs, surplus fund holders prefer debt contracts to sharing the risk of the project through equity participation. In a debt contract, a borrower promises to repay the principal plus an additional sum, the interest, over a stipulated time frame. This, in effect, cuts off the relationship between the project for which funds are needed and its financing, since a debt contract establishes the legal right of a lender to receive more money in the future in exchange for a given amount of principal today – it is an exchange of spot money for more future money – regardless of the outcome of the project undertaken by the investor–entrepreneur. Indeed, if the risks of informational problems and associated monitoring costs are priced into the loan contract, then all risks are shifted to the entrepreneur. One explanation given for the existence of financial intermediaries is that, due to informational problems and associated monitoring costs, it is easier for lenders to delegate their management to a third party, such as a bank, which is thought to be more efficient and experiences economies of scale in collecting information on and monitoring borrowers.

There are ways in which lenders remedy informational problems. For instance, they can ration lending by charging higher interest rates. However, this remedy mechanism also has some adverse consequences: it rations out potentially successful projects as well as risky ones; and it leads to a paradox where, as the lender raises interest rates beyond a certain point, its expected income declines as borrowers reduce or totally interrupt their demand for loans. Second, lenders can ask for collateral – that is, an asset of the borrower whose property rights over that asset are automatically transferred to the lender in the event of default. Third, lenders can include positive and negative clauses in the contract. The former requires, among other actions, a minimum amount of capital to be put up by the entrepreneur. It also requires periodic reports and/or inspections as part of the contract. The negative
clauses prohibit certain actions on the part of the borrower, including restrictions on uses of the borrower’s income, transfer of assets, and prior authorisation from the lender before any financial operation is undertaken. Sometimes, the borrowers themselves try to signal, at times at heavy costs, their trustworthiness, as well as the soundness of the projects they are undertaking. They may, for example, pay for insurance that protects the lenders in case of project failure and/or against the risk of default. Much of the financial structure of modern economies consists of interest-based debt contracts.

Another chief characteristic of modern financial economies is the overwhelming presence of a credit system. The main difference between debt and credit is that, while a debt contract can assume a one-time event between a lender and a borrower, credit is based on a longer-term relationship of trust between the two, as the name implies. (“Credit” is derived from the Latin word credo, meaning “I believe”.) Economic historians suggest that credit, in its modern sense, began with commercial banking (Ferguson, 2008). However, if the term is taken in its broadest meaning – that is, a loan for productive activity – economic historian Sidney Homer (1963, pp. 17–24) traces the origin of credit as far back as 5000 B.C., when “capital and credit became important and provided a main impetus toward human progress”.

In modern times, “credit” refers to an established debt relationship between a financial institution and its borrowing client. If, however, one considers credit as the provision of financial resources to facilitate investment and production, loan agreements based on interest rates are not the only means of credit relations. Such a relationship can be based on the provision of rewards, not on fixed money return on principal, as in a debt contract, but contingent on a project’s outcome – that is, the expected value of the project. In other words, the risks of the project are shared. A debt contract is a fixed nominal obligation with a certain maturity date. A share contract does not have these features. Moreover, a
share is not redeemable. Such a contract does not have the maturity constraint of the debt contract and, provided a market exists, shares obtained through the provision of financing can be sold in case of liquidity need. Thus, contingent payoff, non-redeemability, maturity, liquidity, and risk-sharing are characteristics that distinguish a sharing contract from a debt contract. An advantage of the share contract is that, while monitoring of the investor–entrepreneur and the project may impose costs on the shareholders, because their incomes depend only on the expected value of the project, rather than on the probability of default as in a debt contract, shareholders need not face the risk of moral hazard or adverse selection.

There are other forms of mobilising finance that, presumably, avoid – or at least minimise – informational problems. These include securitisation, leasing and factoring. The first delinks the repayment of a loan from the performance of the borrowing firm by creating an independent entity called a special purpose vehicle (SPV), which buys some assets with predictable revenue streams, such as account receivables, from the firm. Because the risk of the assets owned by the SPV is different from and less than the risk of the original firm as a whole, the SPV can issue new securities that are backed by the underlying assets (called asset-backed securities, or ABS), which it owns. Because of the better credit quality of the underlying assets, the SPV can mobilise funds at a lower cost. Moreover, the SPV can, also at a lower cost, enhance the credit quality of its securities. Such techniques of credit enhancement include tranching of the securities it issues into senior tranches which are repaid as the first claimants of assets and cash flows in case of default, and junior or subordinate tranches which claim residual assets and cash flows after senior tranches are paid. An SPV, or its parent firm, can create a cash reserve account to cover some losses in case of default and/or buy insurance to guarantee payoffs as a means of credit enhancement. Securitisation liquefies existing assets to allow the firm to
acquire other assets and/or pay off expensive debt. In leasing, a firm (lessor) allows another party (lessee) the use of an asset it owns in exchange for a periodic payment. The lessor retains the legal ownership of the asset while it is generating income flow for the lessee. Therefore, the leased asset serves as both collateral and a source of income generation. Factoring allows a firm to sell its accounts receivable to a specialised credit agency or bank at a discount, thereby increasing its own liquidity. While securitisation has mostly relied on debt instruments, such as asset-backed commercial paper (ABCB), an SPV could potentially issue shares allowing greater risk-sharing. As was mentioned, the existence of informational problems in debt markets leads to inefficiencies stemming from attempts by lenders to protect themselves. An important inefficiency is credit rationing and high interest rates, which penalise unfairly low-risk projects or those with potentially high returns. Equity or risk-sharing contracts avoid this inefficiency; however, where and when the level of trust is low, shareholders’ costs of monitoring the entrepreneur and the project and efforts to become informed may not be negligible (Stiglitz, 1987; Stiglitz and Weiss, 1992; Greenwald and Stiglitz, 1990; Hillier and Ibrahimu, 1993; Baltensperger, 1978).

The informational problems that plague the debt market exist between financial institutions and their clients. Among these institutions, commercial banks are highly susceptible to informational problems in their relationships both with their clients and the central bank. In all market-based financial economies, the banks operate on the basis of fractional reserve, where they are required to maintain a fraction of the deposits in readily available liquid form to meet unexpected and sudden large withdrawals. While such reserves are considered as an instrument of controlling informational problems under normal circumstances, in the face of sudden shocks, the failure of one bank potentially becomes contagious for other banks and financial institutions. Such a possibility creates a systemic risk for
the financial system, and for the economy as a whole. Moreover, because the financial system in general and the banking system in particular facilitate and operate, along with the central bank, the country’s payments system, the failure of the financial system can spell disaster for the society as a whole.

To stop the contagion effect of a banking crisis from spreading, governments, through the central bank, guarantee deposits up to some maximum level. This is referred to as “deposit insurance”. Additionally, during normal times, the commercial banks are allowed to use the central bank as their own bank to alleviate short-term liquidity constraints by discounting financial assets at the central bank. Moreover, the central bank can step in and save banks in whose case liquidity problems (due to bad decisions, fraud, mismanagement, etc.) threaten to convert into solvency problems. The central bank “bails out” these institutions to mitigate the risk of contagion and threat to the payments system. The disadvantage of deposit insurance and bailouts is that they give rise to moral hazards, because if a bank or an important financial (or even non-financial) institution is considered “too big to fail”, this encourages undue risk-taking on the part of market players. Through its intervention, the state uses public funds to rescue private lenders unable to service their debts, thus avoiding the collapse of the credit market and threats to the payments system while, at the same time, creating a moral hazard problem that encourages speculative behaviour within the financial system, including within the banking system.

The fractional reserve system molds out of its banks powerful creators and destroyers of credit; for small changes in their monetary base, the banks can expand credit by a multiple amount (Krichene and Mirakhor, 2008). Central banks exploit this ability of the banking system, through monetary policy, to contract or expand credit, depending on the result they wish to achieve; that is, they use the credit-creation ability of the banking system countercyclically. Within a financial
system that operates overwhelmingly with interest-based debt contracts as well as a fractional reserve banking system, there is a mechanism that operates in the opposite direction – that is, it has procyclical effects. Called leveraging, the concept refers to the use of small amounts of equity capital to contract loans that are multiples of the amount of equity. In a modern financial system, generally all financial institutions – banks and non-bank financial institutions – are highly leveraged. Leverage works through the balance sheet of leveraged institutions. To illustrate how this mechanism operates, consider, as a simplified example, the balance sheet of a commercial bank.

Commercial banks and other financial intermediaries are generally very sensitive to changes in anticipated risks and in asset prices, and manage their balance sheets continuously and actively. Their sensitivity to changes in the price of their assets is particularly acute when their net worth reacts to changes in asset prices through an accounting procedure called marked-to-market, meaning that the values of assets on their balance sheet continuously adjust as the market prices of those assets change. The guidance to a leveraged institution for active management of its balance sheet is the concept of value at risk (VaR), defined as the numerical estimate of a financial firm’s worst case loss and is indicated by $V$. Any loss beyond this worst case can only occur with a benchmark probability $p$. Usually, a financial intermediary, a bank, is required to maintain equity capital, $E$, to meet total value at risk for a total asset of $A$ on its balance sheet, such that $E = V \times A$. Leverage is defined as the ratio of total assets to equity capital – that is: $L = \frac{A}{E}$, so that there is an inverse relationship between the leverage ratio and value at risk. The implication is that as the value at risk declines, which happens when asset prices increase during an upswing in financial markets, the leverage ratio increases, meaning that the unit can take on more debt or become more leveraged. For commercial banks, this means higher credit creation.
precisely when credit should contract. Therefore, leverage acts procyclically (Adrian and Shin, 2007).

The leverage ratio of commercial banks in the United States is estimated at around 10 (9.8), much lower than those of the investment banks, estimated at about 20–25 (Greenlaw et. al., 2008). A commercial bank’s assets are composed of securities it holds, including the loans it has extended; its liabilities include the deposits of its clients, as well as its equity capital.

A simplified balance sheet of a commercial bank

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Securities</td>
<td>1. Clients’ deposits</td>
</tr>
<tr>
<td>2. Loans to clients (businesses, households and governments)</td>
<td>2. Equity capital</td>
</tr>
</tbody>
</table>

To simplify further, combine the asset side and assume the balance sheet totals at $100 where the assets are financed by $90 worth of debt (whose price is assumed constant) and $10 worth of equity capital so that the leverage ratio is \( L = \frac{E}{D} = 10 \), which the bank considers as its target leverage ratio; that is, in the case of an increase or decrease in the value of its assets, it adjusts its balance sheet to restore the target leverage ratio. Assume now that the price of assets declines by $1 so that:

\[
\begin{align*}
\text{Assets} & \quad \text{Liabilities} \\
100 & \quad 90 = D \quad \text{asset price declines by}$1 \\
10 = E & \quad \rightarrow \\
100 & \quad 99 \\
\end{align*}
\]

As a result of the decline in the asset price, the leverage ratio changes \( L = \frac{E}{D} = 11 \), requiring a balance sheet adjustment to restore the target leverage ratio. This means that the balance sheet has to contract,
since \( L = \frac{A}{D} \rightarrow 10 = \frac{90}{90} = 1 \). The way the balance sheet is adjusted to restore the target leverage ratio is for \( D \), debt (deposits), to decline by 9 →

**Assets Liabilities**

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>81 = ( D \rightarrow ) asset price reduction of $1 leads to a contraction in credit extended by the intermediary by $9.</td>
</tr>
<tr>
<td>90</td>
<td>90 = ( E )</td>
</tr>
</tbody>
</table>

Under normal circumstances, the intermediaries adjust the balance sheet to restore their target leverage ratio. In times of financial market distress, however, when pessimism leads to panic, the intermediaries increase their target leverage ratio, leading to substantial contraction of their balance sheet leading to reduction in credit. As credit distress in financial markets picks up momentum, a phenomenon known as a credit crunch, itself the end result of deleveraging, develops.

The reverse of this process is at work when asset prices increase. To illustrate, return to the initial balance sheet and assume an increase in asset prices of $1.

**Assets Liabilities**

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 = ( D \rightarrow ) asset price increases by $1</td>
<td>101 = ( D \rightarrow )</td>
</tr>
<tr>
<td>10 = ( E ) \rightarrow 11 = ( E )</td>
<td></td>
</tr>
<tr>
<td>100 100</td>
<td>101 101</td>
</tr>
</tbody>
</table>

As a result, the leverage ratio has changed: \( L = \frac{110}{90} = 9.2 \). To restore the target leverage, 10 = \( \frac{A}{100} \rightarrow A = 110 \); thus the balance expands by adjusting the level of debt (credit):
In times of a rapid increase in asset prices, intermediaries expand their balance sheet by increasing credit by a multiple of the increase in their asset prices. If optimism about a buoyant asset market leads to euphoria, then the intermediaries lower their target leverage ratios or/and move, if possible, some of their assets off balance sheet. The latter lowers the numerator of the leverage ratio, allowing the intermediary to expand credit to restore its target leverage ratio. In a buoyant asset market these developments lead to increased demand for assets, forcing prices to increase beyond the level justified by the fundamentals of the asset itself – that is, a bubble will be created. As Greenlaw et. al. (2008, pp. 29–30) suggest:

“Leverage targeting entails upward-sloping demands and downward-sloping supplies. The perverse nature of the demand and supply curves is even stronger when the leverage of the financial intermediary is pro-cyclical – that is, the possibility of feedback, then the adjustment of leverage and of price changes will reinforce each other in an amplification of the financial cycle. If greater demand for the asset tends to put upward pressure on its price, then there is the potential for feedback in which stronger balance sheets trigger greater demand for the asset, which, in turn, raises the asset’s price and leads to stronger balance sheets. The mechanism works in reverse in downturns. If greater supply of the asset tends to put down word pressure on its price, then weaker balance sheets lead to greater sales of the asset, which depresses the asset’s price
and leads to even weaker balance sheets. The balance sheet perspective gives new insights into the nature of financial contagion in the modern, market-based financial system. Aggregate liquidity can be understood as the rate of growth of aggregate balance sheets. When financial intermediaries’ balance sheets are generally strong, their leverage is too low. The financial intermediaries hold surplus capital, and they will attempt to find ways in which they can employ their surplus capital. In a loose analogy with manufacturing firms, we may see the financial system as having ‘surplus capacity.’ For such surplus capacity to be utilised, the intermediaries must expand their balance sheets. On the liabilities side, they take on more short-term debt. On the asset side, they search for potential borrowers that they can lend to. Aggregate liquidity is intimately tied to how hard the financial intermediaries search for borrowers.”

It should be recalled that fractional reserve banking already gives financial intermediaries the power to create (and destroy) credit (therefore, money) out of thin air (Krishene and Mirakhor, 2008). As shown above, balance sheet adjustment and leverage reinforces this power substantially, to the point of making credit creation out of nothing appear extraordinarily helpful in stimulating economic activity. However, in referring to this phenomenon, Maurice Allais (1987) assets that: “in reality, the ‘miracles’ performed by credit are fundamentally comparable to the ‘miracles’ an association of counterfeiters could perform for its benefit by lending its forged banknotes in return for interest. In both cases, the stimulus to the economy would be the same, and the only difference is who benefits.”

Even if it is assumed that the credit-creating power of the financial intermediaries is beneficial, it is clear that such power also has potentially harmful effects
on economic activity. Because this power operates through interest-based debt contracts and instruments, it dictates the cost of the finance needed for investment; thus, while delinking financing from the underlying productive investment and assets (since it provides money for financing today for more money in the future, regardless of the outcome of the project), the price of the debt contract (interest rate) establishes a benchmark for acceptable rates of return on investment projects in the real sector. Thus, the interest rate on money lent today for more money in the future rules the rate of return to the real sector. Indeed, the financial sector of a modern economy exercises a dominating role over its production sector. Thus, disturbances in the financial sector are automatically transmitted to the production sector, affecting employment and income in the whole economy. Governments concerned with economic growth, employment and the severity of the impact of financial fluctuations on the level of economic activity focus on this aspect of the relationship between the financial and the real sector to manipulate the rate of interest charged by financial intermediaries on credit they extend to entrepreneurs, thus indirectly influencing the rate of return in the real sector. They do this through monetary policy. An active monetary policy places limits on the credit expansion or contraction power of financial intermediaries by making it either easy or difficult for these institutions to access the resources of the central bank or create an incentive for them to become more or less liquid through the sale of government securities, which constitute a significant part of intermediaries’ assets. Thus, to check a downward movement in the level of economic activity, central banks administer easy monetary policy; and to check an upward movement in activity, fearing inflation, they administer tight monetary policies.

In addition, because financial intermediaries, particularly commercial banks, manage the economy’s payments system, and because of the deposit insurance system, which the government makes available to the
intermediaries, the state is justified in regulating and supervising the behaviour and performance of financial intermediaries. Regulation is composed of a set of rules imposed on these institutions to ensure the safety and security of their resources. Supervision ensures that the institutions continuously abide by these rules. One characteristic of regulatory frameworks is that rules are developed in response to past abuses or transgressions not foreseen by previous rules; they are, therefore, ex post frameworks. Given the dynamism of profit-seeking behaviour, financial institutions are constantly in search of ways and means of circumventing the rules of the regulatory framework. This behaviour, referred to as the search for regulatory arbitrage, becomes particularly intense when financial markets become buoyant. Therefore, an efficient regulatory-supervisory framework must dynamically and continuously adjust its constituent rules to new and emerging situations in the financial sector in order to minimise regulatory arbitrage opportunities.

There are times in which monetary policy becomes passive and the regulatory-supervisory framework exercises forbearance in the face of changes in the financial system that could potentially pose serious risks to the economic system. This could occur either in response to political pressure or because of an underlying ideological orientation of policy-makers. Focusing on the latter, policy-makers and regulators may possess an abstract system of thought which is the model upon which they base their own model of behaviour. The policy maker follows a model which is itself grounded on a higher abstract underlying model, i.e., the basic theoretical model that the policy-maker uses to frame policies, is a model of the policy-maker’s model. One such model is derived from a major school of economic thought that holds that markets left to themselves, with minimal government intervention, are capable of producing the best results for the economy as a whole. The intellectual pedigree of this line of thought, known as “neoclassical economics”, dates back to the
classical economists, but its analytic underpinnings were provided in the mid-20th century by the work of Arrow and Debreu (1954) who demonstrated the existence of a general equilibrium for a competitive economy under a set of assumptions that included perfect information, complete markets, no transaction costs and no role for monetary factors. Under these assumptions, equilibrium would prevail in the economy instantaneously, with all resources, including labour, fully employed.

Within the same general framework, two other propositions were developed in the second half of the 20th century that formed the intellectual underpinnings (or the model of the model of the policy-maker) for an ideology of passive monetary policy and regulatory forbearance. The first was the Modigliani–Miller theorem (1958, 1963) which addressed the question of the optimal capital structure of the firm – that is, the best combination of sources of financing investment: equity, debt and internal funds (undistributed profits or retained earnings). The theorem stated that in a perfect capital market, a firm’s value depends on the profitability of the assets that investment generates, and not on how such investments were financed. This implies that a firm should be neutral between sources of financing. This is referred to as the “Modigliani–Miller neutrality theorem”. The second proposition was developed within the neoclassical framework in the 1970s and is referred to as the “efficient markets hypothesis” (Fama, 1970). It suggests essentially that competitive prices contain all the information required for rational economic decision-making. The implication of this hypothesis for the real sector of the economy is that, given that prices include all the needed information, the price mechanism has the capability to allocate resources in such a way as to achieve maximum output. For the financial sector, this hypothesised efficiency simply means that, regardless of time and space considerations, asset prices are always and everywhere established by the demand for and supply of assets about which suppliers and consumers have made rational decisions because the prices of
these assets contain all the information required for such decisions. Therefore, the prevailing asset prices are the correct ones, reflecting the true value of assets. If asset prices change, it is because new information regarding the assets themselves has become available. In other words, the market as a whole has the correct view of asset prices that reflect the underlying fundamental values of the assets. Even if each individual participant in the market (including investors, dealers and lenders) has an inaccurate estimate of the asset values, as a group the market participants will not over- or underestimate the fundamental value of the assets. This would imply no overconfidence, or bubbles, in the market as a whole. Therefore, no matter how wildly asset prices swing, such fluctuations are the market’s response to changes in the underlying fundamental characteristics of the assets themselves, reflecting the scarcity or oversupply of the assets in the market. Asset price bubbles – when market prices deviate substantially from the price justified by the underlying fundamentals – are a natural reaction of the market. Some consider bubbles as beneficial because they stimulate the mobilisation of resources for investment in the assets whose prices are rapidly rising. Moreover, no attempt should be made to stop rising asset prices, because even if bubbles burst, the net effect for society is positive, and because rising bubbles are difficult to predict in the first place (Gross, 2007).

The perfect market paradigm, and the attendant notions of the existence of a competitive equilibrium, neutrality of capital structure (that is, debt-to-equity ratio) and the efficient market hypothesis, all of which were developed between 1950 and 1970, assume perfect rationality on the part of market participants, making decisions with perfect information in a market in which every risk, every commodity and every contingency is fully identifiable (no uncertainty) and representable by a tradable market instrument. That is, it was assumed that for every and all contingencies, a tradable security was available – in other words, the market is complete.
Moreover, the market is assumed to be frictionless, without transactions or information costs. Further, it is fully liquid; trade can take place instantaneously. Financial resources pose no constraints as participants can borrow instantaneously and without limit. Such a market ensures the optimal allocation of resources for maximum output. It is also efficient, as prices adjust instantaneously and appropriately to any new information. This information cannot be predicted ahead of its appearance; it is random and, once it appears, everyone will have it. Market prices adjust quickly to the random information; therefore, prices also change randomly, leaving no room for anyone, including dealers and traders, to earn income from trading in new information. Most importantly, in such a market risks are spread uniformly among all the market participants because insurance against all risks is available to every participant.

It is important to note that, while rigorous mathematical models demonstrate the existence of equilibrium for competitive markets under the above assumptions, the stability of such equilibriums is much more difficult to demonstrate. It turns out that relaxation of the underlying assumptions renders the competitive market equilibrium erratic and unstable. Nevertheless, and despite the unrealism of its assumption, the efficient market is considered as an ideal paradigm toward which all market economies must progress. In this light, volatility and turbulence is caused by shocks that the market experiences because of information imperfection (the informational problems discussed earlier) and/or market incompleteness (not enough security instruments for all contingencies). The ideology of adherence to the efficient market paradigm becomes a model of the model based on which policy is formulated. All financial innovations represent progress toward market completeness, and all volatilities (including the formation and implosion of bubbles) are attempts by the market to adjust to new information. In such a belief system, market intervention may have negative consequences
for the economy. Therefore, the best course of action is a passive policy and regulatory forbearance. Thus, the former Chairman of the Board of the central bank of the United States (the Federal Reserve) considered new financial innovations such as derivatives as a means of “dispersion of risk to those willing and able to bear it” and as instruments that would prevent “cascading failures” (Greenspan, 2002). The implication of the latter is that new instruments make it possible to spread risk around and away from the commercial banks, thereby mitigating the risk of contagion in the banking system (“cascading failures”). The Chairman’s view of using active monetary policy to target bubbles also reflected the underlying ideology. As late as December 12, 2007, in an op-ed page article in the Wall Street Journal, he remarked:

“After more than a half-century observing numerous price bubbles evolve and deflate, I have reluctantly concluded that bubbles cannot be safely defused by monetary policy or other policy initiatives before the speculative fever breaks on its own.”

Given the history of the ideologically based deregulation movement in the United States since the early 1980s, the Chairman’s reference to “other policy initiatives” should not be difficult to understand. The basic idea is that because asset price bubbles are elusive and impossible to spot until they burst, monetary policy should not target rising asset prices in a particular sector since any preemptive action to do so runs the risk of derailing the whole economy (Kohn, 2006). Since the underlying model of the policy-maker’s model is the efficient market paradigm, regulation would be administered with the “light touch” as monetary policy. (For a readable account of the unrealism of the efficient market paradigm, see Cooper, 2008.) It was only after the crisis was in full swing that Mr. Greenspan acknowledged that the efficient market was a flawed

“Among the most astonishing statements to be made by any policymaker in recent years was Alan Greenspan’s admission this autumn that the regime of deregulation he oversaw as Chairman of the Federal Reserve was based on a ‘flaw’: he had overestimated the ability of a free market to self-correct and had missed the self-destructive power of deregulated mortgage lending. The ‘whole intellectual edifice,’ he said, ‘collapsed in the summer of last year.’ What was this ‘intellectual edifice’? As so often with policymakers, you need to tease out their beliefs from their policies. Greenspan must have believed something like the efficient market hypothesis which holds that financial markets always price assets correctly. Given that markets are efficient, they would need only the lightest regulation. Government officials who control the money supply have only one task – to keep prices roughly stable.”

Greenspan, of course, was not alone in thinking that interventions to deflate a bubble may do more harm than good. For example, as late as 29 February 2008, the President and Chief Executive Officer of the Federal Reserve Bank of Chicago, in a speech before the US Monetary Forum, asked: “Should a policymaker deflate a bubble before it becomes problematic?” He then answered: “I am skeptical that we can identify bubbles with enough accuracy and know enough about how to act to say that we wouldn’t have more failures than success…” Furthermore, as former Chairman Greenspan (2004) noted, in order to make sure you burst a bubble, you have to attack it aggressively, because if your attack fails, it just gets bigger. And there are big risks to the real economy.” Similar sentiments were expressed by the President of the European
Central Bank (8 June 2005) in a speech given in Singapore, titled “Asset Price Bubbles and Monetary Policy”.

III. The Crisis, and Explanations of its Causes

It does not seem too long ago that most analysts believed that only developing countries would experience booms and busts, along with the rapid credit expansions and contractions that accompany them. It was believed that the growing sophistication of the financial system, as well as the deeper and greater macroeconomic stability in developed, industrial market economies, had mitigated the risk of emergence of crises. There were, however, warnings from some analysts that, while the likelihood of occurrence of systemic crisis may have been reduced due to better policy-making, the development and growth of financial markets and instruments, better information and the wider dispersion of risk in advanced economies, the potential adverse impact of a crisis, if it occurs, is more severe in these economies than is generally assumed (Gai et. al., 2007; Rajan, 2005) precisely because of rapidly growing complex financial innovations as well as the prolonged period during which major crises were resolved quickly.

The recent crisis began in the sub-prime mortgage market of the US financial system. “Sub-prime” refers to a segment of the market in which mortgages were issued and credit was extended to buyers with a low credit score, brief or no credit history, few or no assets, and poor income-earning prospects. These loans, referred to as “Ninja loans” (no income, no jobs and no assets), were extended in the lending frenzy that was in full swing at least three to four years prior to the start of the crisis. In the period of run-up to the crisis, the US and global economies displayed robust growth, which was expected to continue. Interest rates were low, liquidity was high and growing, financial innovations
were proceeding at a rapid pace (especially in securitisation and structure finance), complacency in the face of growing risk was deepening, and regulation and supervision were receding and weakening. All of this created an incentive structure that encouraged excessive risk-taking in search of higher yields. By March of 2007, the excesses “came home to roost”. High liquidity, low interest rates and easy credit had already created an incentive for home purchases and refinancing of existing mortgages. Prices in the housing market were already increasing, indicating a boom. This provided the primary motivation for the emergence of the sub-prime market, for as long as house prices were increasing, the underlying debt obligation would be continuously validated by the increase in value regardless of the size of the downpayment, the credit record of the buyer and the adequacy of documentation.

The sub-prime borrowers comprised a riskier class, generally with a FICO credit score of less than 640 (a credit score developed by Fair Isaac & Company, ranging between 300 and 850, with a higher score indicating a better chance of repayment), delinquency on some debt repayment in the two years preceding the mortgage application and/or declared bankruptcy in previous years. For this class of borrowers to have access to credit, the mortgages would need to be structured differently and the underwriting standards would have to be substantially weaker. Moreover, real estate developers, appraisers, and insurers and lenders had to cooperate to show a reasonable loan-to-value (LTV) ratio for the mortgage. To encourage borrowers, lenders would offer a 30-year mortgage with a low or zero downpayment and low interest rates for the first two to three years, after which the loan would be reset at double-digit interest rates. Often, mortgage contracts would include a prohibitive cost of early repayment to ensure a steady underlying stream of payment flows. Banks would then package these mortgages (initiated by themselves or real estate developers) and sell them to a special purpose vehicle subsidiary they had established.
In turn, the SPV would securitise them into mortgage-backed securities (MBS) and rearrange them to transform them into newly innovated instruments called collateralised debt obligations (CDO), which would direct the assets into different tranches with different credit ratings, interest rate payments, and terms of priority of repayment.

A given CDO would be composed of hundreds of sub-prime MBS divided into high-rated tranches that paid low interest but with the higher security of being the first to be paid in the event of default, and lower-rated subordinated tranches that paid higher interest rates but with lower security of repayment. The risk appetite of the buyer of these instruments would determine which tranche would be purchased. Further innovation made it possible for an SPV to use these CDOs as an underlying asset to allow another round of packaging them into CDO-squared, which would again be dissected into various tranches and sold. In turn, these CDO-squared could be used for another round of repackaging into CDO-cubed, the next layer of CDOs. On top of these CDOs was another layer of instruments, the CDOs of credit default swaps, which were essentially an insurance written by the originating bank, or a “monoline” insurer that guaranteed the return to “senior” tranches. (For a full explanation of these instruments, see: “Credit Derivatives Explained”, prepared by Lehman Brothers’ Structured Credit Research, available on the Internet.)

This model has been referred to as the “originate and distribute” model. Greater distance was created between the original underlying MBS assets and every new layer of instruments, which were distributed far and wide to all geographically dispersed market participants in the form of structured investment vehicles (SIVs) and conduits. For a fee, banks would establish, usually in a tax and regulatory haven, SPVs, SIVs, or conduits to hold CDOs or MBS on behalf of their clients. The conduits would actively participate in the commercial paper market to borrow short to invest in higher-yielding, longer-term
assets, such as mortgage-based securities or collateral debt obligations.

The basis of structured securitisation was a system of credit arbitrage that needed yet another system of credit rating that would give confidence to the purchasers of these securities that the underlying risks were sufficiently analysed and rated. Credit agencies have the primary responsibility for determining the equity cushion, level of overcapitalisation, and insurance guarantees, as well as other means of credit enhancement, that would permit an instrument to be rated as investment grade. As it turned out, credit agencies were not only approached for their credit rating function, but were also consulted, for a fee, on the design and structuring of a securitisation activity. The SPVs, and other issuers of structured security, would always select the rating agency that granted investment grade to the security at the least possible cost and avoid those agencies that were more cautious. This created an incentive for the credit rating agencies to become more liberal in their risk assessment. As Mark Adels testified before the US Congress on 27 September 2007, the willingness of the credit agencies to allow a deterioration of credit standards – in part to gain business for themselves, in part because of the absence of appropriate historical data on structured securities on which to model the default characteristics of the securities they were rating, and in part because the markets validated less conservative risk assessments – resulted in the erosion of equity cushions being demanded of the structured assets.

Most active participants in these newly developed and complex instruments were unregulated financial institutions, such as hedge funds. The SPVs, themselves established by a parent bank, were off the balance sheet of the bank. This would not only allow the parent bank to liquefy otherwise illiquid (mortgage) assets and expand its balance sheet for greater leverage operation and credit expansion, but would also allow much more room
to manoeuvre for regulatory arbitrage. Moreover, non-
bank financial institutions, such as hedge funds, private
equity funds and SPVs, were mostly free of the
regulatory-supervisory framework which, to an extent,
constrained the banks’ leverage operation. Whereas the
banks’ leverage ratio was about 10, these non-bank
institutions could have leverage ratios as large as 30, as
was the case with the hedge fund Long-Term Capital
Management (LTCM). Increased leverage further
energised an already high level of credit and liquidity in
search of higher yields that could only find outlets in
further risk-taking, particularly in the housing market,
which could go on as long as house prices appreciated.
And, as George Soros (2008) stated in an article in the
New York Review of Books:

“The longer the double-digit rise in house
prices lasted, the more lax the lending
practices became. In the end, people could
borrow 100 percent of inflated house prices
with no money down…. the excesses became
evident after house prices peaked in 2006 and
subprime mortgage lenders began declaring
bankruptcy around March 2007. The problem
reached crisis proportions in August 2007 …
and spread with amazing rapidity to other
markets. Some highly leveraged hedge funds
collapsed and some lightly regulated financial
institutions … had to be acquired by other
institutions in order to survive. Confidence in
the creditworthiness of many financial
institutions was shaken and interbank lending
was disrupted. In quick succession, a variety
of esoteric credit markets – ranging from
collateralised debt obligations (CDOs) to
auction-rated municipal bonds – broke down
one after another. After periods of calm, crisis
episodes recurred in January 2008 … in March
… and then in July…. The deepest fall of all
came in September, caused by the disorderly
bankruptcy of Lehman Brothers in which
holders of commercial paper – for example, short-term, unsecured promissory notes – issued by Leman lost their money. Then the inconceivable occurred: the financial system melted down. A large money market fund that had invested in commercial paper issued by Lehman Brothers ‘broke the buck,’ i.e., its asset value fell below the dollar amount deposited, breaking the implicit promise that deposits in such funds are totally safe and liquid. This started a run on money market funds and the funds stopped buying commercial paper. Since they were the largest buyers, the commercial paper market ceased to function. The issuers of commercial paper were forced to draw down their credit lines, bringing interbank lending to a standstill. Credit spreads – i.e., the risk premium over and above the riskless rate of interest – widened to unprecedented levels and eventually the stock market was also overwhelmed by panic. All this happened in the space of a week.”

Very quickly, the surviving banks and non-bank financial institutions began the process of deleveraging; the credit crunch was in full swing, with what Soros called “the financial system in cardiac arrest”. Conventional explanations of the causes of the crisis pivot on factors such as the frenzied pace of financial innovation, which produced complex financial instruments that only a very few of their purchasers understood, and excessively high liquidity in search of a destination and high yield. This was coupled with lax and inefficient or nonexistent regulatory-supervisory oversight, unfettered growth of financial instruments, such as new derivatives and swaps, plus the fact that high-liquidity leverage within the financial system reached unprecedented heights. As mentioned earlier, high leverage allows expansion of credit to energise asset price increases and the formation of bubbles. The conventional view sees nothing adverse in high leverage
in a dynamic financial system during an upswing. It
does, however, consider deleveraging during the
downswing phase of the asset market as the cause of
financial crises. In this phase of a “leverage-induced
crisis”, financially stressed firms have to liquidate assets
to meet their debt obligations. As many firms have to sell
assets, market prices decline, leading to a decrease in
the value of the collateral of the firms, forcing yet further
liquidation. The firms need liquidity, but those who would
buy assets in normal market conditions are themselves
liquidity constrained or in a state of panic, forcing them
out of the market. Firms needing to meet the demands of
their creditors who are deleveraging must continue to
sell assets in an illiquid market, thus forcing further asset
price declines. The process is referred to as “debt
deflation”, a term coined by Irving Fisher during the
Great Depression. The financial crisis becomes an
economic crisis, which, if unchecked by public policy,
leads ultimately to a depression as investment,
production, income and employment continuously
decline.

The conventional view also considers marked-
to-market accounting rules – which, under normal market
conditions, price assets according to what they would be
worth if they were to be sold at any given moment – as
another factor that enhanced the downward spiral of the
recent crisis. During a crisis, assets cannot be sold
easily, if at all, since there are very few buyers liquid
enough or willing enough to buy them. When there is no
market for an asset, it is said that it has become “toxic”;
and portfolios that include toxic assets are called
“contaminated portfolios”. In such cases, there is no
market to which the asset could be marked, and if a
marked-to-market rule is binding, assets cannot be
valued and cannot serve as collateral for credit that
relieves the liquidity constraint. Markets have a price
discovery function. To perform this function, the markets
require buyers and sellers. When there are only sellers –
but no buyers – in the market for a financial asset that
has some organic relation to other assets in the market,
the market becomes contaminated. If the level of contamination is high enough to affect adversely the price discovery function of the related assets, the first asset is said to have become toxic. The firms, unable to raise credit, will be forced into a “fire sale” of their assets, which often provides insufficient funds for the firm to unwind its positions. In the case of banks and non-bank financial institutions that operate under a minimum equity capital, the fire sale becomes a necessity because these institutions fear falling below their minimum capital requirement. This adds further negative energy to the already existing state of panic in the financial system, enforcing pessimism and the flight to liquidity as a precautionary move, and ultimately feeding and strengthening the ongoing credit crunch, thereby adding to the severity of the crisis. During a crisis, the marked-to-market accounting rule forces the process of asset price determination to become oriented toward liquidity pressures, rather than toward the fundamental values of the assets. This is particularly damaging to highly leveraged institutions as their equity capital base approaches their absolute minimum during the boom phase of a cycle to allow them to expand their credit and acquire high-yield financial assets. In the recent crisis, many banks that had established SPVs had also either guaranteed (explicitly or implicitly) the underlying assets, the assets they originated and sold to their own SPVs, or had established credit lines for the SPVs to buy back the commercial papers the latter issued. These were ways and means by which the banks enhanced the credit of the securities issued by their SPVs. By selling their assets to the SPVs they had established, the banks moved those assets off their balance sheets, which allowed them to expand their balance sheets and, at the same time (or so the banks thought), pass the risks to the buyers of these assets. When the crisis hit, the credit lines and the guarantees forced these institutions to bring these assets back on to their balance sheets, thus contracting them, which led to reduced credit and thus worsened the ongoing credit crunch.
Yet another factor contributing to the emergence of crisis, in the conventional view, has been the risk management model followed by financial institutions. The risk management system the banks and non-bank institutions used, by and large, ignored the possibility of a market crisis, which should have been an important risk in the sub-prime market. It was not recognised that markets are interrelated and that, during a crisis, this interrelationship amplifies. When, during a crisis, one market becomes liquidity constrained, the institutions (or fund managers) are forced to sell assets (or positions) they hold in other markets, producing a decline in asset prices in these other markets and triggering asset sales by highly leveraged institutions (fund managers). Liquidity stress in the first market adversely affects other markets, which may have otherwise been healthy, through contagion. As Soros described, the powerful linkages among markets became obvious during the recent crisis. In a debt-dominated financial system, credit is extended on the basis of an *ex ante*, fixed interest debt contract without an organic relation to the underlying real sector investment. Debt contracts are reinforced by collateral. Asset price increases raise the value of the collateral and the credit quality of the borrower, as well as the credit-creating power of the leveraged institutions. This, in turn, stimulates further demand for financial assets, leading to a cycle of increased asset prices, higher credit, higher borrowing, and greater demand for assets. Asset price declines trigger the reverse movement of the process. “In debt-funded asset markets, price declines beget asset sales that beget more asset declines, morphing into a self-reinforcing positive feedback cycle” (Cooper, 2008).

Since, in the conventional view, market liquidity – especially too little liquidity during the down phase of a financial crisis – plays an important role as a factor contributing to crisis, a digression may be useful. It is argued that, in modern capitalism, financial markets perform two pivotal functions: they provide information
and liquidity. The role of the financial system as a provider of information was discussed earlier and need not be repeated here. Provision of liquidity is considered as a major function of the financial system, far more important than providing information. The efficient market paradigm takes as given that there is sufficient liquidity in the market to allow ready sales and purchases of financial assets, at zero transaction costs, by relatively small market participants who execute their trades without affecting the price – that is, the participants are price-takers, and neither they nor the level of liquidity can move prices. It turns out, however, that liquidity is a crucial factor in allowing financial assets to be bought and sold, at low transaction costs. Most importantly, it is liquidity and the ability to provide it that creates opportunities for profits for financial market participants by influencing the demand for and supply of assets and, therefore, the prices of assets traded. Participants who demand liquidity, and those that supply it, must move the market to the point where the price yields satisfy them. To those who demand liquidity when they offer an asset for sale to liquefy a position, time is the most crucial factor since they are attempting to liquefy one position in order to take a more profitable position once they obtain the liquidity they need to do so.

To the suppliers of liquidity, on the other hand, it is the price that matters. A deviation of the price from what is considered fair (as reflected by the fundamentals of an asset) provides a profit opportunity (particularly in the short run) for when the asset price returns to its fair value. Profit opportunities are embedded in the bid–offer spread. This view of speculation found its first full intellectual support from Milton Friedman (1953) who argued that when market prices of assets diverge from their fundamental value a profitable opportunity is created for speculators to step in and, through trade, force prices back to their fundamental values. The market intermediaries whose roles are crucial in the provision of liquidity in the financial market are: (i) market makers and specialists who operate on the exchange
floors of the markets and whose job is to move the prices in accordance with asset sellers who need liquidity; (ii) broker–dealers in markets where there are no central exchanges, such as markets for corporate bonds and MBS; and (iii) hedge funds and other leveraged institutions that earn profits through speculation. The first two market players focus on short-term trade, while the speculators take a long-term view in their trading. The conventional view holds that, far from playing a parasitic role, these intermediaries perform a vital function in the financial system and the economy as a whole by providing liquidity to the market. Through this service, they make financial resources available to market participants, manage risk, make expert judgements regarding the differential between the market price of a financial asset and its fundamental value, and, through execution of trades, assist in the emergence of equilibrium in the market. For all these services, they deserve remuneration. In the recent crisis, however, the conventional view holds that market conditions were such as to create incentives for these players to take excessive risks, which proved disastrous when the down phase of asset market prices began.

To summarise, the conventional view holds that the emergence of the crisis was due to a number of factors, including: extraordinarily high liquidity; rapid pace of financial engineering which innovated complex, opaque and difficult-to-understand financial instruments way ahead of the market; informational problems caused by lack of transparency in asset market prices, particularly in the market for structured credit instruments; outdated, lax or absent regulatory-supervisory oversight, which encouraged excessive risk-taking; faulty risk management and accounting models; and the emergence of an incentive structure that created a complicit coalition composed of financial institutions, real estate developers and appraisers, insurance companies and credit-rating agencies whose actions led to a deliberate underestimation and underpricing of risk. The conventional view suggests that mitigating the risk
of emergence of similar financial crises in the future requires: (i) “efficient”, rather than “more”, regulations and supervision that avoids excesses in risk-taking; (ii) prudentially higher capital requirement or collateralisation to reduce leveraging; (iii) improving risk management to focus on market and systemic risk, rather than the risk to individual institutions; (iv) designing and adopting an improved system of due diligence by financial institutions and financial assets and instruments; (v) designing and adopting an improved system of liquidity management by institutions participating in the financial market that would mitigate liquidity risk – that is, the chance that a holder of an asset may not be able to sell it quickly; (vi) designing and establishing a last resort liquidity provider that can buy assets from liquidity-stressed market participants and hold them in order to infuse liquidity into the market until it returns to normalcy; (vii) creating an incentive system for banks to bring on board and consolidate off-balance sheet items to which they have extended (explicit or implicit) guarantees and credit lines; and (viii) revising the marked-to-market accounting rule to avoid enforcing and accelerating the downward pressure on asset prices in times of liquidity stress. (For details of the conventional view of the crisis and its causes, see various documents of the International Monetary Fund (IMF) 2006–2007, including World Economic Outlook, Global Financial Stability Report and Finance and Development; and various publications (including speeches by governors and the chairmen) of the Federal Reserve System, the Bank for International Settlements and the European Central Bank.)

There is an alternative view that sees crises as endogenous, endemic to financial capitalism. More specifically, instability and crises are inherent in a financial system where the predominant mode of financing is interest-based debt contracts and credit. According to economic historians, from its very beginning, profit-seeking for the purpose of wealth accumulation has characterised capitalism and
distinguished it from other economic systems. At the heart of the financial wealth-creating institutional framework of financial capitalism is a system of debt–credit creation centred in the commercial banking sub-system that places a multiple of the deposits of many depositors at the disposal of borrowers of money capital. In essence, every deposit expands into a multiple of itself to become loans (credit) in the hands of the borrowers (Krichene and Mirakhor, 2008). In the process, the banks make money off the interest rates they charge on the money they lend for more money in the future. Part of the borrowed money finances new investments, part finances purchases of financial assets, and the rest finances current consumption. The only part of the created credit that helps in the expansion of the economy is the part that is borrowed to finance investment, which provides the future flow of income to validate the debt obligation that made the financing of investment possible.

It is worth noting that, contrary to the common view, consumption today validates the debt obligations incurred for financing of past investments. A fall in consumption represents a failure to validate payment obligations made in the past. For validation of past and present obligations, investment and consumption must grow. Debt and credit, however, grow or decline for reasons different from those that cause investment and production to grow; there is no direct connection between credit expansion and investment in the real sector (for production). There is an indirect connection between them through the interest rate mechanism. The interest rate is set by the financial institutions (banks and non-bank financial intermediaries) for loaning money. These institutions are only tangentially interested in the real rate of returns to an investment project, and that only as a signal that payment obligations will be validated. To ensure that this will be the case, lenders ask for collateral. As such a financial system becomes more sophisticated and more dynamic, ways and means will be sought through innovations to financialise as
many commodities and real assets as possible. The process of financialisation, it is argued, expands access to finance, thus reducing the force of financial constraints. The financialisation process transforms illiquid real economic resources into asset classes that are traded in various asset markets. If and when a financial sector is dominated by interest rate-based debt contracts, the financialisation process creates more and more debt as it expands throughout the economy, converting equity in real assets into debt. This was the case in the early stages of the housing boom in the United States, where excess liquidity and low interest rates created an incentive for homeowners to cash out equities built up in their homes through refinancing. The cashed-out equity supported a consumption boom.

It is noted that the securitisation that has been ongoing over the past three decades is part of the general process of financialisation. While it is quite conceivable that financialisation (or securitisation) could be just as easily equity based, the dominant force of socio-political factors and the power structure have created incentives for debt-based financialisation. The result has been rapidly growing corporate debt-to-equity and household debt-to-income ratios; acceleration of dominance of the financial sector relative to the real sector; income transfer from the real sector to the financial sector; deterioration of income distribution and increased income inequality; and changes in the orientation of the US economy from a saving–investment–production–export orientation to one of borrowing–debt–consumption–import (Palley, 2007).

This reorientation has transformed economic productive activities so that they resemble participation in the activities of a casino, as Keynes remarked, or in those of a racetrack, both of which use real resources but produce no real output, no productive investment (Hirschleifer, 1971). Such an economy produces “rolling bubbles” in financialised assets. As one bubble bursts, finance moves to another. Such has been the case over
the past three decades, as bubbles were created and then imploded in the emerging market debt, dotcom, real estate and commodities markets. In none of these bubbles investments in real productive activities were the primary objective of debt and credit expansion. Expectations of higher prices for the financial assets attracted participants in droves, creating bubbles. That this would happen was analytically demonstrated as early as the 1980s. For example, Flood and Garber (1980) demonstrated that rational individuals participate in asset price bubbles if they have expectations of rising asset prices. Growth in liquidity, low interest rates, higher leverage and rapidly expanding credit, combined with regulatory-supervisory forbearance and passivity, accelerate the emergence and growth of bubbles.

Growth in the volume of debt in the United States over the past three decades has been the defining characteristic of the process of financialisation. Palley (2007) estimates that the total credit-to-GDP ratio grew from 140% to 328.6% of GDP between 1973 and 2005, while mortgage debt-to-GDP grew from 48.7% to 97.5% over the same period. Household debt-to-GDP grew from 45.2% in 1973 to 94% in 2005. Palley also presents data demonstrating that financialisation not only has led to the rapid expansion of debt, but has also contributed to adverse “changes in the functional distribution of income, wage stagnation, and increased income inequality” (Palley, 2007). In Section II, reference was made to the Modigliani–Miller theorem of neutrality of the corporate financial structure (represented by the debt-to-equity ratio), which argued that how a firm finances its investment is irrelevant to its value. However, within the US financial system, there is an incentive structure in place that biases the financial structure in favour of debt. The US tax code gives interest payments a more favourable treatment than dividends and profits. Moreover, corporations can use debt and interest payments to reduce other claims (workers) on its income stream. Additionally, the rate of return on equity capital of the firm increases through debt financing, which allows it to expand leverage.
Corporations as well as consumers “have been encouraged to adopt a cult of debt finance. To reinforce the process of progressive reliance on debt finance, asset price inflation provides consumers and firms with collateral to support debt-financial spending. Borrowing is also supported by steady financial innovation that ensures a flow of new financial products, allowing leverage and widening the range of assets that can be collateralised.” (Palley, 2007). A theoretical foundation of justification of financialisation has been the work of the Arrow–Debreu framework. In this model, financial assets represent contingent claims, not those resulting from ex ante, fixed-return debt contracts. That is, claims are validated if a future contingent state materialises. These contingent claims result from financing of investment in the real sector. When a firm finances an investment in the real sector by issuing an ex ante, fixed-return debt, the result is not a state-contingent claim – that is, it does not depend on the outcome of the investment project for which financing was obtained. The debt contract requires repayment on the date specified by the contract regardless of the outcome of the project. In effect, the financing bears no relation to the real investment. While it is true that the investor undertaking the project must compare the rate of return to the project to the rate of interest specified by the debt contract, the financing itself (that is, the lender, amounts borrowed, the debt contract, and the rate of interest charged) is, in effect, decoupled from the real sector investment. Thus, the rate of interest specified in a debt contract, which is essentially a promise to pay more money in the future for an amount of spot money, dominates the rate of return to investment in the real sector without the financing itself having an organic relation to that investment.

While in the early stages of the growth of the debt-dominated financial system there is a tenuous relationship between financing and real sector investment through the conduit of an entrepreneur’s comparison of the expected rate of return to the investment project and the rate of interest, as
financialisation proceeds and debt securitisation grows in sophistication, the relationship becomes progressively less important. As the financial sector grows to dominate the real sector, layer upon layer of securitisation thins the connection between the two to the point where an inverted pyramid of debt is supported by a very narrow base of real sector output and assets. The overwhelming dominance of the financial sector over the real sector can be discerned by noting that the ratio of global financial assets to the world’s annual output of goods and services grew from 109% in 1980 to 316% in 2005. Similarly, while the total world GDP was about US$48 billion in 2006, the value of global financial assets in the same year was US$140 billion (nearly three times as much). As of 2007, the global liquidity market was estimated to be 12.5 times as large as the global GDP. Financial derivatives constituted 80% of this liquidity. As Lim (2008, p. 13) suggests, a point has been reached “where what happens in the financial markets affects, or perhaps dictates, what happens in the real economy. It is the case of the tail wagging the dog.”

The alternative explanations of financial crises view them as internally generated instability episodes that inevitably arise from the basic debt–credit–interest rate relations. Where a financial system is dominated by interest-based debt and credit contracts, a fundamental “conflict between guaranteeing return of capital while also putting that capital at risk is a key channel through which financial instability can be, and recently has been generated” (Cooper, 2008). Fractional reserve banking and its close relatives in the form of money market funds and other financial innovations operated by highly leveraged institutions ensure that the credit (and debt) creation process amplifies manifold. This takes place through the mechanism of a money-credit multiplier within the fractional reserve banking system, and through leverage ratios within the banking system as well as other highly leveraged financial institutions. Credit multiplies during the upswing phase of a financial cycle, when financial asset market price bubbles
emerge; and is rapidly destroyed during the downswing phase, when bubbles burst. “In money markets, as with most debt markets, the way to earn highest rates of interest is to make loans for the longest possible periods to the lowest quality, least reliable investors. The pressure for high money market yields, therefore, encourages fund managers toward a high-risk lending strategy. But this strategy runs into direct conflict with the money market fund’s commitment to give back all of the investor’s money, plus interest earned, without the risk of losses” (Cooper, 2008). What is true of money market funds is much more forcefully true of other highly leveraged financial institutions such as hedge funds, SPVs, SIVs and others. The paradox of a debt-dominated, dynamic financial system is that, as it innovates layer upon layer of debt-based financial instruments, the real (productive) sector of the economy (the part of the economy that, ultimately, has to generate income streams to validate all the repayments of financial capital, plus the interest rates stipulated ex ante in debt contracts) shrinks in size relative to the financial markets. As noted above, Lim (2008) estimated that the global size of the financial assets in 2007 was 12.5 times that of global GDP! That such an inverted pyramid would eventually collapse under its own weight seems inevitable. Additionally, debt contracts essentially pit ex ante, fixed and guaranteed payment commitments against an expected but uncertain future income stream from the underlying investment in the real sector to validate the payment of principal and interest stipulated in the debt contract. It is no wonder, then, that Keynes referred to such a system as a “casino”.

The warning signs of such an eventual implosion had been around long before the recent crisis. Indeed, five years before the event, it was observed (Mirakhor, 2002) that:

“While the financial innovations of the 1990s in the conventional system have led to mobilisation of financial resources in
astronomical proportions, they have also led to equally impressive growth of debt contracts and instruments. According to the latest reports, there are now US$32 trillion of sovereign and corporate bonds alone. Compare this (plus all other forms of debt, including consumer debt in industrial countries) to the production and capital base of the global economy, and one observes an inverted pyramid of huge debt piled up on a narrow production base that is supposed to generate income flows that are to serve this debt. In short, this growth in debt has nearly severed the relationship between finance and production. Analysts are now worried about a ‘debt bubble’. For each dollar worth of production there are thousands of dollars of debt claims."

The succeeding five years made this picture far more ominous as debt grew further, its growth rate dwarfing the rate of expansion of the global production base. For example, by 2007, credit default swaps alone had grown in size to more than US$50 trillion, as compared to the total US GDP of US$14 trillion.

The view that a financial system dominated by credit and debt contracts is prone to instability and eventual collapse has been around since the 19th century. But its most respected intellectual pedigree dates to the years of the Great Depression when the view found forceful expression in the writings of eminent economists, such as Fisher and Simons (see Krichene and Mirakhor, 2008) in the United States and Keynes (1930, 1936 in the United Kingdom). The recognition that the fractional reserve banking system, in which credit multiplier and leverage ratio mechanisms were operative, was the source of credit instability led American economists, including Fisher and the Chicago Group (including Henry Simons, Frank Knight and other members of the economics faculty in the University of
Chicago), to propose the reform of the US banking system, to require banks to maintain reserves equal to 100% of their deposits. While the proposal was not enacted into law (by some accounts, due to the political pressure of the banking lobby; see Phillips, 1995) at the time, it has resurfaced from time to time in a variety of formations, such as “narrow banking”, “collateralised banking” and others (see Bossone, 2002; Phillips, 1992a and 1992b, 1995; Konstas, 2006; Minsky, 1994; Pierce, 1991; Spong, 1993; Scott, 1998; Garcia et. al., n.d; Wallace, 1996; Kobayakawa and Nakamura, 1999). The more recent discussion of 100% reserve banking proposals focuses on the moral hazard and costs of deposit insurance and lender of last resort functions, which central banks have to establish to cover deposits in the fractional reserve banking system in case of defaults, and on the advantage that the proposed reform will have in imposing discipline on credit creation within the system.

Whereas these American economists saw the fractional reserve banking system and its power of credit creation as a source of financial instability, Keynes saw another (deeper) “villain of the piece”: the role of interest and the rentiers who demanded it. Keynes saw this role as being so detrimental to the economy as to call for steps toward the “euthanasia of the rentier” in his book, *The General Theory of Employment, Interest and Money* (1936). The issue of interest rate as rent was important enough to Keynes that he devoted a good part of the book to it (particularly in Chapters 12, 17 and 23, which are also among the most neglected chapters). A reputable economist whom no less than Joan Robinson called “a Keynesian” considers the notion of the “euthanasia of the rentier” among the 11 most important elements constituting the “classical Keynesian position”, as opposed to what Robinson called a “bastard Keynesian” position. By the latter, it is meant a system of economic thinking and policy that deviates profoundly from the views held by Keynes but which, nevertheless, its adherents refer to as “Keynesian” (Turgeon, 1996).
As important as the detrimental impact of an *ex ante*, fixed rate of interest on debt (particularly that which did not lead to investment) was to Keynes, it is puzzling and enigmatic, yet intriguing, that the economics profession has all but ignored this elemental concept of Keynesian thought. Even when a "Keynesian" does refer to this concept, the coverage tends to be superficial, historically as well as analytically. For example, consider Turgeon’s brief reference to this concept:

“Some wag has defined an economist as someone who has seen something work in practice and then proceeds to make it work in theory. In some respects, this may have applied to Keynes, who was certainly aware of the tremendous economic miracle of Adolf Hitler in reducing unemployment from over 30 percent when he took office in 1933 to 1 percent by 1936, the year in which the German edition of the *General Theory* appeared. One of the features of Hitler’s economy that also shows up in the *General Theory* is the neutralisation of monetary policy. Like other populists, Hitler had a special aversion to interest as a form of income since he considered this return to capital to be ‘parasitic.’ Keynes in the *General Theory* has only a few references to monetary policy and, in one instance, he prescribed a lowering of interest rates in what we would later refer to as an overheated economy. He also prescribed the ‘euthanasia of the rentier’ or the eventual elimination of real interest in a mature capitalist economy. Subsequently, Keynes’s disciple and early biographer, Sir Roy Hurrad, would also suggest the abandonment of interest as a category of income to pacify postwar socialists in Great Britain.” (Turgeon, 1996)
The above remarks are analytically and historically wrong and a clear misrepresentation if intended to suggest, as the first sentence in the above quotation implies, that Keynes saw that Hitler’s repugnance to interest “worked in practice” and then proceeded “to make it work in theory”. The remarks also seem to make a vague suggestion of a timing coincidence between Hitler’s views on interest rates and those of Keynes. If this reading of the remarks is correct, it is historically off the mark by a wide margin. Historical evidence suggests that Keynes formulated his position on interest at least in the second half of the 1920s and presented it in his Treatise on Money, published in 1930, before Hitler appeared on the scene. Moreover, as will be seen, his views on interest and the rentier were the subject of a heated debate in the Economic Journal between 1931 and 1932, in which Keynes himself took part. Additionally, Keynes was not alone among British scholars in holding such views; nor were expressions of it confined to the 20th century (see, for example, Ferguson, 2008). The idea of a rentier–interest rate relation – that there is a class in market economies that lives on interest income and finds advantage in holding liquid assets rather than risking their holdings in employment-creating investment and would part with their liquidity only if they can loan it in the form of iron-clad debt contracts that guarantee full repayment of the principal and interest – is an essential element in the explanation Keynes provides as to why a market economy left to itself will not create full employment. The rentier class will require a guaranteed ex ante specified rate of interest before parting with their liquid resources, which the entrepreneurs need to finance their employment-creating investment. It is then the entrepreneur who will have to determine if the expected rate of return from his investment (marginal efficiency of capital) is high enough to allow him to meet the demands of the rentier. There are two ways to remedy the situation: elimination of the rate of interest, or “euthanasia of the rentier”. Keynes chose the latter method, advocating a public policy that would socialise
capital investment combined with a low-interest policy. He believed that it was the scarcity of capital that gave the financial capital owners the power to extract a rent (interest) as a price for giving up their preference to hold their assets in liquid form.

A careful reading of the Treatise on Money and The General Theory makes clear that Keynes did not believe that there is either a theoretical explanation or an economic justification for the existence of an \textit{ex ante}, fixed (or even variable) interest rate payment which (along with the principal) was guaranteed by a debt contract. His own liquidity preference theory explained why the rentiers were demanding an \textit{ex ante} return on the money they would lend, rather than providing a justification (theoretic or economic) of the existence of such a rate. Nor has any theory emerged subsequently that can explain or justify such a rate as an integral part of a coherent economic model (see Iqbal and Mirakhor, 1987). In fact, Keynes himself develops a theoretical edifice in terms of the concept of “own rate of interest” to suggest that any commodity in the spot market will have a rate of return in its term in the futures market, which may be zero, positive or negative, but most certainly not \textit{fixed ex ante} as was the case of a rate of interest on money demanded by the rentier. Moreover, throughout his writings, Keynes emphasised that the future is uncertain – in the Knightian sense; an uncertainty not reducible to risk to be insured. In such a world, rates of return to assets could not be known \textit{ex ante}. Whereas the marginal efficiency of capital (the rate of return on real sector investment) is determined within the real sector of the economy, an \textit{ex ante}, fixed rate of interest on money is “determined by psychological and institutional conditions” (Keynes, 1936, pp. 202, 217).

While Keynes did not advocate outright elimination of the rate of interest by direct government intervention, he did consider that it is quite possible for the interest rate to converge to zero “within a single generation”, making “capital goods so abundant that the
marginal efficiency of capital is zero, this may be the most sensible way to gradually getting rid of many of the objectionable features of capitalism. For a little reflection will show what enormous social changes would result from a gradual disappearance of a rate of return on accumulated wealth.... Though the rentier would disappear, there would still be room, nevertheless, for enterprise and skill in the estimation of prospective yield about which opinions could differ. For the above relates primarily to the pure rate of interest apart from any allowance for risk and the like, and not to the gross yield of assets, including the return in respect of risk” (Keynes, 1936, p. 221). Clearly, Keynes made a distinction between return to entrepreneurial and financial resources willing to take risk, which would be contingent on the outcome of the real sector investment undertaken by the entrepreneur and financed by the risk-taker, the surplus-fund holder, and an ex ante-determined rate of interest on money required by a rentier to part with liquidity. The latter he considered as “the villain of the piece”.

The Treatise on Money was published in 1930 when economic deterioration was picking up momentum to become the Great Depression. Shortly after the book’s publication, a scholar by the name of H. Somerville published a short article in the December 1931 issue of the Economic Journal which prompted a debate on issues covered in the Treatise, including saving, investment, taxes, debt, credit, interest rate, usury, and scholastic thought. Somerville argued that “one of the unexpected consequences of Mr. Keynes is a vindication of the Canonist attitude to interest and usury”. By “Canonist”, Somerville was referring to a system of thought expounded by scholastic scholars (also called “Churchmen” or “Schoolmen”) of the Middle Ages who, in combining faith and reason, explained many religious precepts of Christianity. The most famous among these scholars was St. Thomas Aquinas. In his book History of Economic Analysis (1954), Joseph Schumpeter, after discussing the economic thoughts of the Greeks and
Romans, claimed that an intellectual gap of 500 years separated the contribution of Greco-Roman thought to economics and scholastic thought, which he referred to as “the Great Gap”. Schumpeter states that, as far as economics is concerned, “we may safely leap over 500 years to the epoch of St. Thomas Aquinas (1225–74) whose *Summa Theologica* is in the history of thought what the South-Western Spire of the Cathedral of Chartres is in the history of architecture” (pp. 73–74). In a paper presented in 1983, Mirakhor offered evidence that such a gap (great or otherwise) never existed. He traced the thoughts of scholastic scholars, to the Islamic world and its scholars, whose writings and thoughts were transmitted to the scholars of the Middle Ages via a variety of channels (see Mirakhor, 1983 reproduced in Al-Hassani and Mirakhor, 2003 for a fuller discussion). Among the economic thoughts of scholastics was the prohibition of interest charges on money loaned. By claiming that the position which Keynes took in his *Treatise* on the question of saving without investing, but demanding a reward anyway, vindicated the scholastic view on prohibition of interest, Somerville triggered the heated debate in the *Economic Journal* from December 1931 to March 1932 when the Great Depression was in full swing. Particularly controversial was Somerville’s claim that:

“It is an inescapable conclusion from the Keynesian analysis that interest is the villain of the economic piece. Not that Mr. Keynes suggests the possibility of abolition of interest. According to his theory, interest could be too low and might require to be raised in the general interest: this would be in circumstances when prices were advancing too rapidly. Leaving aside for a moment the old question whether interest is necessary to evince saving, the only use that Mr. Keynes can see for interest is as a depressant in times of over-activity.
The orthodox doctrine has related interest closely to profits as if the two progressed or declined together. Mr. Keynes shows them as antagonists. Interest upon money is simply an added cost upon capital goods and therefore a deduction from profit and a burden upon enterprise. Socialist theory assails interest even more destructively than does Mr. Keynes, but socialist theory also assails profit, whereas Mr. Keynes salutes profit as the engine that drives the car of progress.”
(Somerville, 1931, pp. 647–648)

It is worth noting that by the “orthodox doctrine”, Somerville does not mean the “Canonists” but the acceptable body of economic thought of his own day. The above remarks indicate that he is not a socialist, a popular school of thought of the time. He makes his own position clear:

“The Cardinal point made by Mr. Keynes is the distinction between Saving and Investment, between the saving of money and its conversion into capital-goods. Saving without investment is not a service to production, and the saving of money does not by itself cause any conservation of products, but their waste, or their disposal at lower prices and the slowing down of productive activity…. From the point of view of general economy it might be better to spend on production than on consumption, but certainly the worst thing is the sterile saving of money. The saving is sterile from the stand point of community, though it may be profitable to the individual who, apart from any interest earnings, may watch his idle money grow in purchasing power through the fall in general prices. In such conditions, it is a matter to be deplored that holders of money are able to get interest by bank deposits. Interest only encourages
socially wasteful saving and discourages socially desirable investment. To this extent, therefore, and in such conditions, interest is anti-social. Mr. Keynes urges the lowering of interest, even the abolition of interest on bank deposits, as proper policy to be pursued when trade is below par.” (Somerville, 1931, p. 647)

Somerville also reiterated that the position of the Canonists on the impermissibility of interest as usury did not extend to profits earned as a result of risk-taking. This is the point of convergence of the views of the Canonists on interests and profits with those of Keynes. He stated that “the Canonists never quarreled with payments for use of capital, they raised no objection to true profit, the reward of risk, ability and enterprise, but they disputed the identification of the lending of money with the investment of capital and denied the justice of interest as a reward for saving without investment…. The Canonist principle was that sharing in trade risks made an investor a partner, a co-owner of capital, not simply a money lender, and gave a title to profit” (Somerville, 1931, p. 648). To support his understanding of the scholastic position on interest and profit, Somerville referred to Ashley’s *Economic History*, one of the leading textbooks of the time on the subject (see Mirakhor, 2003), particularly Book II of Ashley’s text, in which it is asserted that “until the beginning of the sixteenth century it was the constant teaching of the Canonists that to bargain for a fixed reward, or dividend, upon the capital invested, whatever the fortunes of business might be, made the contract usurious”. Somerville concluded his article by suggesting:

“There may be reasons for thinking that the world will go back to the early Canonist doctrine. The classical argument that interest is necessary to evoke saving wears a different aspect when we appreciate that saving does not necessarily mean investment. The saving of money may actually diminish investment,
and interest is deterrent to investment. If we could ensure, as the Canonists tried to do, that saving should be rewarded only when it was also investment in capital-goods, we should have gone far to stop the master-evil that Mr. Keynes has revealed to us, of saving exceeding investment.” (Somerville, 1931, p. 649)

Somerville’s potent argument that, based on his understanding of Keynes’s main thesis in the Treatise, the cause of underemployment and inherent instability in a money-capitalist system is the fact that in such a system, where debt contracts dominate investment financing, without government intervention, there is no way to guarantee equality of saving and investment, and that the interest rate is the mechanism responsible for this state of affairs, brought an immediate response. Three of the four papers published in the March 1932 issue of The Economic Journal (by Edwin Cannan, B.P. Adarkar and B.K. Sandwell) under the heading: “Notes and Memoranda: Saving and Usury: A Symposium”, were highly critical of Somerville, accusing him of: misunderstanding, either the Canonists’ position or that of Keynes, or both; of “misconceptions” of saving and investment; and of “hostility to personal ownership” (pp. 123–135). The fourth, and last, paper was by Keynes himself (Keynes, 1932, pp. 135–137) who was supportive of Somerville, both in the latter’s understanding and presentation of Keynes’s arguments in the Treatise as well as in the position of the Canonists. Keynes’s article primarily focuses on Edwin Cannan’s article, saying:

“… Prof. Cannan agrees with Mr. Somerville that if saving is conceived as mere refraining from expenditure, or if it is conceived as saving up money, ‘the case against interest as a consequence of saving is black.’ But, he continues, ‘the answer is that interest is not, in fact, obtained as a consequence of saving in
either of these two senses. No one gets a 
penny of interest in consequence of merely 
refraining from expenditure; no one gets a 
penny of interest in consequence of having 
merely saved up money.’ I wish I could agree 
with him in attributing this natural justice to the 
economic system, but I am sure it is not so. 
Prof. Cannan has, I think, overlooked a vital 
aspect of the argument in my Treatise on 
Money wherein it differs from what I was 
brought up to believe and continued to believe 
until recently.

The point is this. The answer to the 
question whether there is an increment of 
wealth corresponding to the saving of an 
individual seldom depends, as Prof. Cannan 
claims, on what he does with the money which 
represents that part of income which he 
refrains from spending on current 
consumption. In particular, the answer does 
not depend, as Prof. Cannan seems to 
suggest, on whether he ‘hoards’ the money by 
increasing his cash or uses it to buy a security 
or some other capital asset. He may use his 
savings to buy a bond, and yet there may be 
no increment of capital wealth coming into 
existence as a result of his saving. I have 
argued in my Treatise that the causes which 
determine the increment of capital wealth are 
only contingently and indirectly connected with 
those which determine the amount of 
individual savings. If an increment of saving by 
an individual is not accompanied by an 
increment of new investment – and, in the 
absence of deliberate management by the 
Central Bank or the Government, it will be 
nothing but a lucky accident if it is … then it necessarily causes diminished receipts, 
disappointment and losses to the other party,
and the outlet for savings of A will be found in financing losses of B.

Thus, when an individual saves, his savings must be balanced by the creation of either of an asset or a debt (or a loss paid for by an asset changing hands). But, as a rule, it lies entirely outside the power of the individual saver to determine which it is to be, and whether the result, or rather the accompaniment, of his saving is to be an asset or a debt. What he has done is to make possible the creation of an asset without a rise in price-level. But failing a simultaneous increment of new investment, either by good management or by a lucky accident, then his act of saving will cause an equal loss to someone else; a debt will be created or an asset will change hands, but there will be no increment of wealth.

Does Prof. Cannan hold that if an individual increases his bank-deposit by ‘saving up’ money out of income, there necessarily results an increment of wealth to the community? If so, this is a view with which I have tried to join issue in my Treatise on Money; if not, he has failed to meet the point.

Now when an act of saving merely results, however unintentionally, in a loss to someone else, it is of an anti-social tendency, and the subsequent payment of interest to the saver – for a pace Prof. Cannan, debts have to pay interest just as much as assets – is a burden which, if accumulated with time, may become insupportable.

That is why, without contesting anything in Mr. Adarkar’s note, I nevertheless agree with Mr. Somerville that it is this social evil, to the possibility and theoretical explanation of which
I drew attention in my *Treatise*, which probably lay behind the doctrine of the Canonists.” (Keynes, 1932, pp. 135–136)

The implication is clear: it is interest-debt that is a “social evil” and the “villain of the piece” of the explanation for unemployment, fragility and, ultimately, the inherent instability of debt-based financial capitalism. Without the institution of interest, unspent income would find its way to investment in exchange for a reward contingent on the outcome of the real sector investment, for which savings provided the financing. Emboldened by the support his views received from Keynes, Somerville produced another note for the June 1932 issue of the *Economic Journal*, in which he mounted a well-argued attack on the institution of interest. The centrepiece of this note focused on the fact that the economics profession had not produced a satisfactory theory or explanation of why an *ex ante*, fixed interest had to be paid on debt. One by one, Somerville refuted the existing explanations and restated the Canonists’ position on usury, again demonstrating the convergence of this view with that of Keynes with considerably more clarity:

“Canonist legislation prohibited all interest ... but freely allowed profits even to sleeping partners.... Interest was forbidden, while profit was allowed, because interest arose simply from a loan of money and profit from an investment of capital.... In striking at interest on money loans, the Canonists were striking at saving-without-investment apart from that which consists in the simple hoarding of currency. At the same time, they were positively encouraging, both by their prohibition of interest and their allowance of profit, investment in the sense of the production of capital goods....
Money-lending without investment was recognised practically and theoretically as a social evil in the fourteenth century. Mr. Keynes’s ‘Saving without Investment’ is correctly translated as money-lending without investment, it being understood that savings deposits are regarded as lending.... The lending of money necessarily involves a debt, and it never by itself constitutes an asset.... The great support which Mr. Keynes gives to the Canonists, I take to be this: his strong distinction between saving and investment shows that it is theoretically wrong to treat money as representative capital and lending as investment. Interest is the price paid for the use of money, it is not the yield of capital.” (Somerville, 1932, pp. 322–323)

Keynes went on to expound his ideas on issues raised in the Economic Journal Symposium in his General Theory with greater clarity. Toward the end of his book (Chapter 23), he makes the following remarks regarding the Canonists’ position on usury:

“I was brought up to believe that the attitude of the Medieval Church to the rate of interest was inherently absurd, and that the subtle discussions aimed at distinguishing the return on money-loans from the return to active investment were merely Jesuitical attempts to find a practical escape from a foolish theory. But I now read these discussions as an honest intellectual effort to keep separate what the classical theory has inextricably confused together, namely the rate of interest and the marginal efficiency of capital. For it now seems clear that the disquisitions of the schoolmen were directed towards the elucidation of a formula which should allow the schedule of the marginal efficiency of capital to be high, whilst
These remarks seem to validate Somerville’s conception of convergence between the views of scholastic scholars on interest and profit and those of Keynes. Be that as it may, in the General Theory Keynes makes the point forcefully that the interest rate, through its role in creating a wedge between saving and investment, creates conditions under which the system becomes inherently unstable. In his concluding remarks (Chapter 24), Keynes suggests:

“I feel sure that the demand for capital is strictly limited in the sense that it would not be difficult to increase the stock of capital up to a point where its marginal efficiency had fallen to a very low figure. This would not mean that the use of capital instruments would cost nothing, but only that the return from them would have to cover little more than their exhaustion by wastage and obsolescence, together with some margin to cover risk and the exercise of skill and judgment. Now, though this state of affairs would be quite compatible with some measure of individualism, yet it would mean the euthanasia of the rentier, and, consequently, the euthanasia of the cumulative oppressive power of capitalist to exploit the scarcity value of capital. Interest today rewards no genuine sacrifice, any more than does the rent of land. But whilst there may be intrinsic reasons for the scarcity of land, there are no intrinsic reasons for the scarcity.” (Keynes, 1936, p. 376)

He viewed the “euthanasia of the rentier” as a gradual process, one that “will need no revolution”. This will happen when “a somewhat comprehensive socialisation of investment will prove the only means of
securing an approximation to full employment; though this need not exclude all manner of compromises and devices by which public authority will co-operate with private initiative" (Keynes, 1936, p. 378). Thus, there would be “an increase in the volume of capital until it ceases to be scarce, so that the functionless investor will no longer receive a bonus…” (Keynes, 1936, p. 376).

Keynes believed that financial capitalism, left to its own devices, is inherently unstable. At the core of this belief are the real phenomena of saving and investment processes. Coming basically from two different sub-sectors of the real economy, consumer and business, their coordinated behaviour is subject to uncertainty. That even under the best of circumstances their equality is not always assured, is at the core of his explanation of the inherent instability of the system. The existence of a financial system dominated by ex ante, fixed interest-based debt contracts makes achieving sustained full employment equilibrium difficult if not impossible. The essence of Keynes’s prescription for achieving stability and full employment in such an economy was socialisation of capital investment, which would allow diminished scarcity of capital and the eventual “euthanasia of the rentier", to ensure that all savings would be channelled to productive employment-creating investment. In Chapter 24 of his General Theory, Keynes argued that interest rewards no genuine sacrifice, and that its compounding, which leads to wealth accumulation at an accelerated pace, ensures that wealth and income distribution in society is tilted toward the rentier. The end result is unemployment, poverty and deprivation. Thus, he argued that the chief “evils” of modern financial capitalism were its inability to provide full employment and its strong tendency to generate an arbitrary and inequitable distribution of income and wealth. Through his call for the “euthanasia of the rentier", he linked both these “evils” to their underlying mutual cause: the institution of interest.
One of the most perceptive, productive and brilliant followers of Keynes was Hyman Minsky who pushed forward the frontiers of “the classical Keynesian” (as opposed to “bastard Keynesian”) thought to produce valuable insights into the workings of the financial capitalist system. As did Keynes before him, Minsky considered such a system inherently unstable, holding that if the financial system is dominated by \textit{ex ante}, fixed interest-based debt contracts, the structure itself becomes a source of amplification of disturbances. His major contribution is known as the “financial instability hypothesis” (Krichene and Mirakhor, 2008; Mirakhor, 1985). This hypothesis contains two main propositions; the first states that there are two financing structures: one promotes stability, the other instability. The second proposition states that, in the financial system of money capitalism, stability is not sustainable because, during prosperity, stability contains the seeds of its own instability. Minsky referred to this proposition as “stability is destabilising”. A simple rendition of the first proposition is to say that the more a financial structure (debt-to-equity ratio) tilts toward debt, the greater the fragility of the system. This is what happens, according to Minsky, to the financial structure of the economy through time. In a period of prosperity, there are large payoffs to borrowing used to finance activity in areas and sectors of the economy with profitable opportunities. Initially, businesses are conservative and finance their activities through equity finance and/or from their own internal funds. And, if they borrow, they do so only if their future income streams are sufficient for them to meet the payment commitment stream (principal and interest) over the lifetime of the contracted debt. Minsky calls this “hedge finance”. The financial system dominated by this type of financing (mostly equity and internal funds, with some debt commitments that are validated comfortably by an underlying income stream) is consistent with stability. However, as profit opportunities intensify during prosperity, there is higher reward to borrowing as enterprises take greater risks. Thus, more and more firms and other participants tilt their financial structure
toward debt and increased leverage, and become "speculative" units. That is in response to what they see as profitable opportunities and in expectations of exploiting them. These firms overwhelm the financial structure with debt, to the point where their income stream becomes insufficient to pay the principal. Although speculative units pay interest when due, they have to rollover the principal at maturity. According to Minsky, matters do not rest here. Firms continue to borrow to the point where their financial structure is made of debt commitments that can be validated by more borrowing to pay both principal and interest. Minsky referred to these units as "ponzi units" and to their financing as "ponzi finance". He believed that during the prosperity phase of a business cycle, capitalist economies tend to become progressively more fragile as their financial structure changes from hedge (little or no debt) to speculative and ponzi finance (Minsky, 1986).

The pivotal element of Minsky’s financial instability hypothesis is debt. So important is this element that Minsky considered his hypothesis as a “theory of the impact of debt on system behavior”. There are two forces that push debt financing to higher and higher levels in the upward phase of the cycle. First, market participants borrow more and more because asset price increases validate their expectations, which undergirded their increased borrowing in the first place. Moreover, as prices increase, the value of their collateral increases and with it their creditworthiness, allowing them to borrow more. Second, banks and other highly leveraged financial institutions expand credit and push lending in two ways: (i) as prices of assets increase, their balance sheets expand, allowing them to extend more credit; and (ii) they find new ways and means of credit expansion through financial innovation. These financial intermediaries are, after all, what Minsky called “merchants of debt”. They are constantly searching for ways and means of expanding their balance sheets. Thus, the debt structure continues to be extended
throughout the financial system and beyond to the whole economy.

Minsky was one of the most astute students and observers of capitalism. He saw it as a dynamic system that is constantly evolving. Within it there are a number of dialectical processes and feedback loops at work that made issues of instability, unfair distribution and unemployment structural problems of the system. In this he was following Keynes. And, like Keynes, he thought the dialectic forces within the system would lead it into disaster if the system were left to its own devices. There were ways in which the system could be stabilised – for instance, Minsky believed the solution was a “Big Government”; big enough so that its expenditure would be a stabilising force and serve as an “employer of last resort” and a “Big Central Bank” to serve as an effective lender of last resort. Additionally, Minsky called for a dynamic regulatory system that would be constantly ahead of the curve to minimise the likelihood of regulatory arbitrage (see Minsky, 1986).

In the aftermath of the recent crisis, many have found Minsky’s diagnoses of past crises, and his explanations of potential turbulences that he foresaw ahead, to be insightful and enlightening, as he had warned of growing fragility in the system, debt buildup in the household and business sectors, and the adverse potential of securitisation, debt globalisation and wrong-headed government policies. He had placed great stress on the adverse impact of ongoing ideologically-based deregulation that began in the early 1980s. A number of his colleagues, former students, and followers carried on this tradition after his death in 1996, analysing the unfolding events in the financial sector using his financial instability hypothesis. In a number of papers, published between 1996 and when the crisis hit, these scholars warned of an impending disaster as they observed the growing debt and fragility in the system. (For these papers, see various working papers and policy briefs of The Levy Economics Institute of Brad College, available
on line.) They saw the phenomenon of bubbles forming, inflating and bursting not as an isolated incidence and as due to external factors, but as “rolling bubbles” – symptoms of growing financial fragility that would eventually lead to full instability. (Soros, too, had seen each bubble as being part of a “super bubble” of growing debt and credit – Soros, 2008.) Minsky had observed the growing fragility of the US financial system since 1966, where a bubble’s boom and bust in one asset market was followed by the formation and implosion of another bubble in a different asset market. These booms and busts saw major players fail and have to be bailed out by the government. They included the emerging market debt crisis, and the LTCM, dotcom, housing and commodities bubbles. Just as financialisation would help to create one asset market after another, expanding liquidity and credit in search of yield would create one bubble after another. In the aftermath of the recent crisis, a number of Minsky’s colleagues and students have analysed these events using his framework (see Whalen, 2008; Wray, 2008a and 2008b).

IV. The Islamic Financial System and Lessons of the Recent Crisis

Over the past few decades, a consensus emerged that expansion of credit and debt is detrimental to the stability of developing economies. For example, the IMF advised its developing country members that in order to mitigate the risk of instability, such as occurred in the emerging markets in 1997, they should: (i) avoid debt-creating flows; (ii) rely mostly on foreign direct investment as external financing; (iii) if they must borrow, ensure that their external debt is never larger than 25% of GDP and that their debt obligations are not bunched toward the short end of maturities; (iv) ensure that their economy is producing large enough primary surpluses to meet their debt obligations; (v) ensure that their sovereign bonds incorporate clauses (such as majority action, initiation and engagement clauses) that make debt workouts and
restricturings easier – that is, to ensure that there exist better risk-sharing mechanisms associated with their debt obligations to avoid moral hazard; (vi) ensure that domestic corporations have transparent balance sheets, follow marked-to-market accounting, and have financial structures that are biased more heavily toward equity and internal funding and are not heavily leveraged; and (vii) ensure that their domestic financial institutions are regulated and supervised efficiently, are not highly leveraged, follow prudent credit policy and are highly transparent. The IMF strongly prescribed periodic audit of the soundness and stability of the financial sector of its members’ economies. While a majority of the members complied, some of its major shareholders, notably the United States, repeatedly refused such diagnostic audits. The financial crisis has proved that major countries are in need of IMF financial policy prescription as much as, if not more than, developing countries.

Be that as it may, it is now clear that financial systems dominated by interest-based debt contracts are prone to financial fragility and instability. Nevertheless, other than Keynes and his most ardent followers, few, if any, notable economists have proposed outright elimination of ex ante, fixed interest-based debt contracts in practice. Theoretically, however, the workings of such debtless systems have been investigated in the form of theoretically modelled systems, such as pure “stock market” economies and “cash-in-advance” systems, to analyse their implications. One of the earliest analytically elegant models of a stock market economy was developed by Lloyd Metzler (1951), who investigated the economic implications of an economic system in which private wealth is in only two forms: “money (including demand deposits) and common stock, and that all common stock involves appropriately the same degree of risk”. Metzler further assumed that the central bank is legally authorised to buy and sell the common stock held by the owners of private equity and that this common stock constitutes the only non-
monetary asset of the banking system. The Metzler model further assumed a closed economy, a fixed labour supply, and that all means of production, other than labour, were produced at constant returns to scale. With these assumptions, Metzler defined a “rate of interest” as “nothing more than the yield of the stock, and this yield, in turn, is the ratio of the income earned by the stock to its market price” (Metzler, 1951). It is clear that this “rate of interest” is not the same as the customary \textit{ex ante}, fixed interest rate on borrowed money stipulated in a debt contract. For one thing, Metzler’s “interest rate” is determined by the earnings and price of the stock, clearly an \textit{ex post} concept. Second, unlike the customary \textit{ex ante}-fixed interest rate debt-contract where total interest paid is tied to the amount of money loaned, Metzler’s interest rate was based on the performance of the stock in the market and the earnings from the investment activity that issued the stock; again, an \textit{ex post}, rather than an \textit{ex ante} concept. Metzler then proceeded to drive the equilibrium conditions for such a system and investigate its stability characteristics.

Sensing that Metzler’s model was a reasonable first approximation of an Islamic financial structure, since it assumed away the existence of debt instruments, Mohsin Khan (1986) constructed a simple version of the Metzler model to demonstrate its equilibrium. He showed that the system produces a saddle point – that is, the equilibrium is stable. Khan then suggests that:

“Based on principles of equity participation, Islamic banking may well prove to be better suited to adjusting to shocks that result in banking crises and disruption of the payment mechanism of the country. In an equity-based system that excludes predetermined interest rates and does not guarantee the nominal value of deposits, shocks to asset positions are immediately absorbed by changes in the value of shares (deposits) held by the public in the bank[,] therefore, the real values of assets
and liabilities of banks in such a system would be equal at all points in time.” (Khan, 1986)

Khan and Mirakhor (1989) structured a different model to show that in such a system, monetary policy is not impaired. Both these models were closed-economy models. Zaidi and Mirakhor (1988) constructed an open-economy general equilibrium model to investigate the implications of operation of an Islamic financial system, particularly the effects on the economy’s capacity to adjust to disturbances, and on international capital flows. It concluded that monetary policy is effective for stabilisation purposes and that disturbances to asset positions are absorbed efficiently in an Islamic financial system. These models focused on the financial sector without directly involving the real sector of the economy. Mirakhor (1990) constructed a model that incorporated some of the characteristics of the above models and features of some well-known models, as well as insights by researchers who had extended Metzler’s model in new directions. The latter research work had demonstrated that a fully equity-based system has desirable features that improve the shock-absorption adjustment capacity of the economy; that is, such a system is more stable than a debt-based system, as it adjusts rapidly to shocks (see, for example, Shane, 1984; Cole, 1988). Like the Metzler–Khan model, the model in Mirakhors’s paper assumed only two assets: money and shares; but unlike the former model, where the rate of return was given as the ratio of the return to stock and its market price, Mirakhor’s model derived a rate of return from the real sector, thus providing an interactive process between the financial sector and the real sector which provided the rate of return to the former sector. The equilibrium conditions in the short and long run for both closed- and open-economy models were derived. The paper then investigated the stability of the equilibrium and the process of adjustment of the system to shocks.
In summary, standard economic analysis can demonstrate the stability of an Islamic financial system and its resilience to shocks, at least in theory. The intuition behind this conclusion is that, unlike a debt-based system, there is a one-to-one mapping of the financial structure onto the underlying real sector assets. There is neither the problem of mismatching household savings and finance for investment, nor of mismatching maturities. Risks of loss are shared between the surplus fund holder and the entrepreneur. There is no opportunity to expand credit and leverage beyond what can be supported by the real sector output (Krichene and Mirakhor, 2008). It has already been remarked that, as conventional analysis has demonstrated, the informational problems that are characteristic of debt-based system do not exist in the case of share contracts; thus, risk-sharing, equity-based financing is more efficient. All of this has been demonstrated by standard analysis in a conventional system. But each financial system operates within an institutional framework that facilitates its efficient operations. If the elements of that framework are impaired, the efficiency of the system is adversely impacted. The stability characteristics demonstrated by the above-mentioned theoretical papers are generated within a conventional institutional framework. Even so, the superior stability characteristics of a non-interest-based equity system, as compared to an interest-based debt system, have been demonstrated. Scholars, some of whose works were briefly reviewed above, have argued that it is the predominance of debt that is responsible for the inherent instability of a debt-based market economy. The point is that it is possible – at least theoretically – to envision a non-interest-based financial structure within a conventional system. What makes for a truly Islamic system, however, is not only the prohibition of interest (although this is an important element of the system), but the institutional framework within which the system has to operate. Without its institutional underpinnings – a platform of behavioural rules – Islamic finance becomes indistinguishable from a conventional system, with
financing being provided through sharing, rather than debt contracts.

The institutional framework – what Douglass North calls the institutional scaffolding – that Islam provides for the operation of its economic and financial system strengthens considerably the stability of the system. The elements of this “scaffolding” come directly from the Qur’an, with additional explication and operationalisation from the Sunnah of the Messenger. This framework includes, inter alia, sanctity of contract (explicit and implicit); property rights; trust; rules of behaviour in governance; existence of markets; rules regarding the allocation, production and distribution of resources, income and wealth; rules governing the behaviour of market participants; and rules regarding post-market distribution. (See Mirakhor, 2007, Iqbal and Mirakhor, 2007; and Askari et. al., 2009, for a more detailed discussion of the elements of the institutional framework of Islam.) There can be little doubt that with such a strongly rules-based framework based on faithfulness to contracts and a strong prohibition against taking interest, lying, cheating and other fraudulent activities, the financial system of Islam would be transparent, efficient and informationally trouble-free. For example, consider the implications of full operationalisation of only one element of Islam’s institutional scaffolding – that is, the first verse of Chapter 5 of the Qur’an: the rule of faithfully abiding by the terms and conditions stipulated by a contract to which one is a party. Its full implications are nothing short of astounding. No one would need fear that a contract would not be performed as a result of lying, cheating, fraud or negligence. Imagine the efficiency gains in such a system; there would be no monitoring costs, no risk of moral hazard or adverse selection, and a minimal cost of transaction in contracting.

While there are many whose behaviour corresponds to the rules specified by the institutional framework of Islam, the scaffolding in its entirety is not
yet fully in place in Muslim countries. Until such time as Islamic rules of behaviour are fully operationalised, the available institutional framework in Muslim societies will have to be organised to ensure that non-correspondence of actual behaviours in society with those expected by Islam does not create damaging distortions, dislocations and exploitations that harm the welfare of the members of the society. It is in this context that the recent crisis holds valuable lessons for Islamic finance. As discussed earlier, there are two explanations of how the crisis developed. One view holds that crises are endemic to a financial system that is debt denominated, and where contracts are made on the basis of money now for money later without much consideration of relations to activities in the real sector. Such a system creates an incentive structure for the rapid expansion of credit and debt, through leverage, and the emergence of asset price bubbles. As market players in such a system, financial institutions are “merchants of debt”. As rising asset market prices validate round after round of rising profit expectations, these “merchants of debt” are encouraged to increase leverage, through financial innovation, expand credit and, in turn, encourage their clients to tilt their financial structure ever more toward debt, thus making it more and more fragile. When asset market prices reach a limit of increase and interest rates rise for fear of inflation, some of these debt obligations are not validated. Because of the complexity and interrelated nature of financial markets, failures to validate a few large debt payment obligations create a rapid contagion effect as more and more portfolios in the market become contaminated. As Soros has observed, bubbles have a built-in asymmetric nature. It takes a lot longer for them to emerge, but they implode in short order. As the authorities bail out major participants from the disastrous impact of one asset market bubble implosion, creating lower interest rates to minimise the impact on the real economy and on other asset markets, and higher liquidity, innovations create new asset classes to which liquidity in search of yield migrates. The result is a “rolling bubble”, except that each bubble is
larger and packs a greater force than the one before it. Soros suggests the emergence of an immensely sized “super bubble” within which smaller bubbles are created and then busted. Such was the case in the US financial system as bubbles rolled from emerging markets debt, to the dotcom, real estate and commodities markets. At each turn, bailouts, low interest rates, inflow of funds from emerging markets and financial innovations created a powerful source of funds in search of yield, which poured into new asset classes. The first view holds that these results were predictable and predicted, as financial fragility (greater reliance on debt) was reinforced. It holds that the “ideologically-based aversion to regulation” that prompted the “deregulation revolution” of the last quarter-century in the United States, as well as in other advanced economies, created an incentive structure for accelerating the rapidly paced, debt-based financial innovations which, in turn, were powerful stimulants to the emergence of consumer–business debt binges.

The more conventional view, on the other hand, tends to downplay the idea of “super bubbles” and “rolling bubbles”, as well as the thesis of the inherent instability of a debt-based financial system. It treats bubbles in isolation, each having different reasons for its emergence, but due mostly to the “irrational exuberance” of market players. It further holds that bubbles cannot be predicted; therefore, attempts by policy to target rising prices in a given asset market may have an adverse impact on economic growth. Moreover, this view argues that risk-taking is an essential element of the dynamism of a market economy. Therefore, too much regulation that thwarts risk-taking will harm economic growth. Financial innovations are to be welcomed not only because they are necessary to the dynamics of risk-taking but also because they help in the process of completion of markets and, thus, increase efficiency. Asset price increases are a natural response of the market to investment opportunities, and bubbles are the result of an overly exuberant response of market participants to profit opportunities. (Sooner or later,
market forces will dampen the over-exuberance and bubbles will disappear.) As indicated earlier, the former Chairman of the US Federal Reserve was one policy-maker who espoused such a view, although he recently admitted that it was based on a faulty theoretical model.

V. Conclusions

The major lesson of the recent crisis for Islamic finance, especially at this juncture in its evolution, is the need for the design and development of a comprehensive and dynamic regulatory-prudential-supervisory framework, uniquely and properly designed for an Islamic financial system. Such a framework will satisfy the requirements of any existing regulatory framework anywhere in the world, and go beyond them to ensure the stability of the system. Theory has demonstrated the stability of an equity-based, risk-sharing financial system. Moreover, theory can also demonstrate, easily and comfortably, that the institutional framework (rules of behaviour) of Islam, within which its financial system must operate, reinforces to a high degree the stability and efficiency of the financial system. In practice, however, and as long as the institutional framework is not fully in place, an Islamic financial system can fall victim to the same adversely designed incentive structure as the conventional system, particularly because Islamic finance is presently operating in an institutional framework which is basically that of the conventional system. Mere declaration of prohibition of ex ante, fixed interest-based debt contracts by fiat, but without any effort at implementation of the supporting institutional framework, will not accrue the benefits of properly structured Islamic finance to the population; indeed, it may do harm by creating a sense of complacency in the society. Creating a non-interest-based system side-by-side with a conventional system within an institutional framework that is basically designed to support the latter system has benefits and costs. Which of the two dominates is an empirical question. However, from pure
intuition, the benefits would seem to overwhelm the costs in the long run when all possible benefits are considered, as it allows an orderly evolution of Islamic finance.

In either case, a properly designed regulatory-prudential-supervisory framework seems essential to the orderly development and evolution of Islamic finance. Such a framework will have to be uniquely designed to distinguish it from that of a conventional system, in that it has to structure its incentive system to include both negative elements (don'ts) that specify prohibited behaviour, and positive elements (dos) that encourage and enforce recommended behaviour. Such a framework will have to be comprehensive, covering all transactions, and all financial instruments and institutions operating in the system, without exception. One of the most damaging elements of the US regulatory system was that a larger segment of the financial markets and institutions had little or no regulatory-prudential-supervisory oversight. Moreover, the regulatory framework was fragmented. In an Islamic financial system there are no interest-based debt contracts; financial innovation can proceed only as it relates to equity and trade-based transactions (both of which areas provide ample opportunity for financial engineering). Nevertheless, the risk of inappropriate, ostensibly trade- or equity-based, but actually debt-like, instruments that are highly collateralised to enhance their credit rating cannot be minimised. The soundness and appropriateness of some of the reverse-engineered financial instruments innovated in recent years have been questionable (Iqbal and Mirakhor, 2007; Mirakhor and Zaidi, 2007; Chapra, 2007; Hassan and Lewis, 2007; Siddiqi, 2007). It is the responsibility of an appropriately designed regulatory oversight to ensure that such risks are mitigated. Moreover, such a framework will have to be unified under one regulatory umbrella. The artificial segmentation of the financial markets into money and capital markets for regulatory purposes was one of the most damaging aspects of the
regulatory framework in the United States, in which the regulatory authority was segmented between various agencies – the Federal Reserve, the Comptroller of Currency, and the Securities and Exchange Commission – the states and their regulatory apparatus, and the commodities market with its regulatory-supervisory rules. Given the rapidly paced financial innovation, the line of demarcation between money markets, capital markets and all sorts of financial asset markets is blurred. In an Islamic financial system, the segmentation of regulatory authority is even less logical because of the nature of a system that promotes nearly a one-to-one interaction between real sector and finance activities.

The properly designed regulatory oversight for Islamic finance will have to be uniform – that is, its standards must apply uniformly to all Islamic financial institutions, transactions and instruments everywhere, globally across all jurisdictions; the reason being that the systemic risk of a failure of one Islamic financial institution or instrument is far greater for Islamic finance than for its conventional counterpart because of the potentially magnified reputational damage for the entire system, particularly at this juncture of the evolution of Islamic finance. Because profit-seeking motivation currently drives financial-sector innovations, there are strong incentives to create instruments and ways and means of regulatory arbitrage. An appropriately designed regulatory framework for Islamic finance will have to have sufficient built-in flexibility and dynamism to allow it to stay ahead of the innovation curve, in order to minimise the risk of regulatory arbitrage. An appropriately designed, unified, uniform-comprehensive framework, administered globally, requires a universal mandate to enforce its rules and standards everywhere. Such a mandate requires all governments in countries where Islamic finance is in operation to extend such a mandate to the unified global regulatory agency for the development of oversight standards of Islamic financial institutions, transactions and instruments. Given the design and implementation of such a framework, even
governments of non-Islamic countries where Islamic finance operates should have no difficulty with such a mandate because they, too, are primarily concerned with the safety and security of all financial institutions operating within their borders. No less an ardent supporter of free markets than Lionel Robbins once remarked that “the pursuit of self-interest, unrestrained by suitable institutions, carries no guarantee of anything but chaos” (Robbins, 1952). A regulatory framework is the most important of “suitable institutions”. While important in the conventional system, it is crucial for an Islamic financial system at this nascent stage of its development because of the heavy costs of materialisation of reputational and systemic risks of the failure of a few, or even one, Islamic financial institution or instrument, given the present sensitivities.

Consequently, the most important lesson of the recent crisis for Islamic finance is an urgent need for the design, development and implementation of a comprehensive, unified, uniform, global and dynamic regulatory-prudential-supervisory framework. Such a framework of standards needs a unified administrative agency to uniformly and globally enforce the rules of the framework across jurisdictional boundaries. This requires the agency to have a legislatively based mandate to enforce implementation of its standards and oversight rules. Clearly, this is a most serious challenge, but one that must be met because, at the present stage of development, Islamic finance is embedded in an institutional framework designed to support a conventional financial system and in which innovations replicate instruments designed for a conventional system through “reverse engineering”. In such a system, there is no assurance that financial bubbles and their boom and bust cycles can be avoided. Because the materialisation of such events has significant reputational risk for Islamic finance, to a degree that may abort its further development, the design and implementation of an appropriate regulatory framework, as described above, is urgently needed.
REFERENCES


Askari, Hossein, Zamir Iqbal and Abbas Mirakhor (2009), New Issues in Islamic Finance and Economics: Progress and Challenges (Singapore: John Wiley & Sons Asia).


Friedman, Milton (1953), Essays in Positive Economics (Chicago: The University of Chicago Press).


Homer, Sidney (1963), A History of Interest Rates (New Brunswick, New Jersey, USA, Rutgers University Press).


Iqbal, Zamir and Abbas Mirakhor (2007), An Introduction to Islamic Finance: Theory and Practice (Singapore: John Wiley & Sons Asia).


Turgeon, Lynn (1996), Bastard Keynesianism (Westport, CT: Praeger).


