A·U·G·U·R

Challenges for Europe in the world, 2030

WP2 Financial markets

Systemic Risk

John Eatwell

Queens' College, Cambridge

Second draft. 11th November 2011

Introduction

In their important post-crisis study of financial regulation, Brunnermeier, Crockett, Goodhart, Persaud and Shin (2009) define systemic risk in terms of its regulatory implications

> "Macro prudential regulation concerns itself with factors that affect the stability of the system as a whole. ..the nature of regulation applied to an individual institution depends crucially on how "systemic" its activities are. This is related .. to its size, degree of leverage and interconnectedness."

This characterisation of systemic risk has resulted in a growing body of research and of policy proposals related to the three key components, including debates over "too big to fail", systemically important financial institutions (*sifis*), leverage collars and liquidity regulation, and increased interest in the linkages between financial institutions (Adrian and Brunnermeier, 2009; Haldane, 2009; Allen and Babus, 2009; Allen, Babus and Corletti, 2010).

A common feature of most of this analysis is that it tends to be *ex post* in that it takes the structure of markets and the inter-connections between institutions as given. The structures and linkages do not merge from an *ex ante* model of the financial system. Moreover, systemic risk arises not from the macro-system as such, but from the micro-economic linkages between the agents.

The starting point in this paper is rather different in that the focus is on the evolution of the structure of the financial system, and the consequential macro-systemic risk that is generated.

The argument developed below derives from two sources: congestion theory as first developed by Pigou (Pigou, 1912; see also Vickrey, 1969, and Walters, 1961) and network theory as developed for the analysis of congestion in transportation and telephone systems (see Kelly, 1991). An analogy is drawn between congestion and the "density" of the financial system, defined as the gross value of financial transactions relative to the underlying event being financed.

The changing nature of financial markets

One of the most striking phenomena of the years following the liberalisation of financial markets in the 1970s, has been the very rapid growth of financial assets relative to the growth of GDP. In very broad terms, the assets of the banks, for example, have growth at an average rate of 15% since 1978. Given that the world gdp has grown (in nominal terms) at a little over 5.8% per annum over the same period, the excess growth of 9.2% per year suggests that the banks' balance sheets are now around 20 times greater, relative to the given underlying gdp, than was the case 33 years ago[1]. In the absence of comprehensive transactions data, the growth of financial assets is a proxy for the growth of transactions[2]. Since deposits are not likely to rise at a rate much faster than growth of gdp, the increase in the size of financial balance

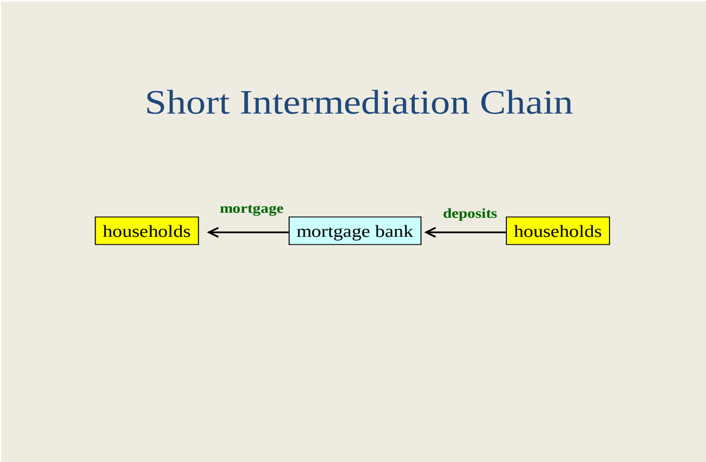
[1] One of the most dramatic increases has been in the relationship between forex transactions and the volume of international trade and long-term investment. In 1971 the ratio between transaction and trade was between 2 and 3. Today it is estimated to be in excess of 80.

[2] The growth of transactions is likely to be higher than the growth of assets, not least because many of the transactions create assets in the non-bank financial sector (including "shadow banks").

sheets must be due to lending between financial institutions and other wholesale activities.

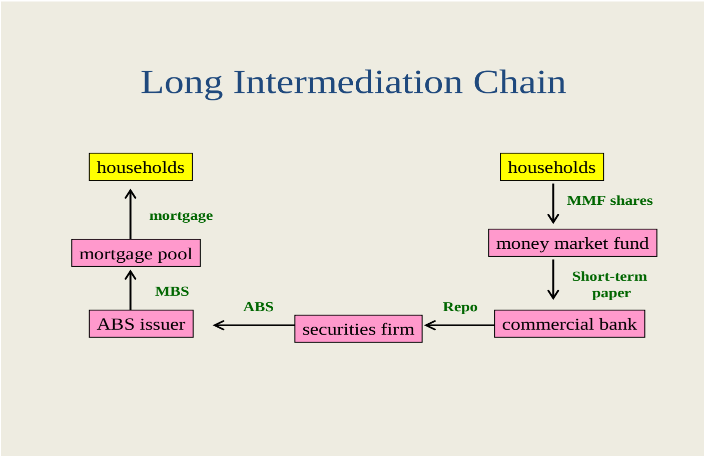
A simple example of what has happened can be seen in the market for domestic mortgages (see Shin, 2010). In the 1960s the financing of mortgages involved households depositing funds in mortgage banks that were then lent on to other households to enable them to buy houses (Figure 1). Today this transaction is likely to pass through a long chain of investments, from the household purchase of money market funds, to short-term loans to the bank, which expands funding through repo transactions with a securities firm that in turn purchased securities from a provider of asset backed securities, that were in turn assembled from a mortgage pool created by lending to home-buying households (Figure 2). Indeed, the illustration is probably a rather a short, uni-directional chain.

Figure 1



If the value of the underlying mortgage transaction is \$100k then there are \$200k worth of financial transactions associated with the intermediated transfer of funds from the depositing household to the home-buying household. *Gross* assets of \$200k are created – \$100k of assets in the form of a bank deposit on household's balance sheet, and a \$100k mortgage on the bank's balance sheet.

Figure 2



The major features of the long intermediation chain are that it is longer (obviously), that the assets created along the chain are likely to be of shorter duration, and the gross value of assets in relation to the underlying is far greater in the longer chain than in the shorter chain. The lengthening of the chain was due to straightforward economic motivation. Changing the length of the chain derived from a combination of increased liquidity, raised return and reduced micro-risk to the trading counterparties.

The lengthening in the chains of intermediation has been reflected in turn in the composition of banks' balance sheets. Figure 3 displays the typical UK bank balance sheet in 1964. The liabilities side of the balance sheet is almost entirely deposits from firms and households. The asset side consists of around half liquid funds (cash, acceptances, treasury bills) and around half loans and advances. By 2007 residents' deposits comprise only a fifth of liabilities, the remainder consisting predominantly of loans from other banks, repos and other short-term commercial paper (Figure 4). The asset side of the balance sheet is similarly transformed, being now dominated by *marketable* securities, repos, and proprietary trading (investments). The other two significant changes are the internationalisation of the balance sheet, and the fact that bank assets in 2007 amounted to 497% of gdp, compared with 34% of gdp in 1964.

Figure 3

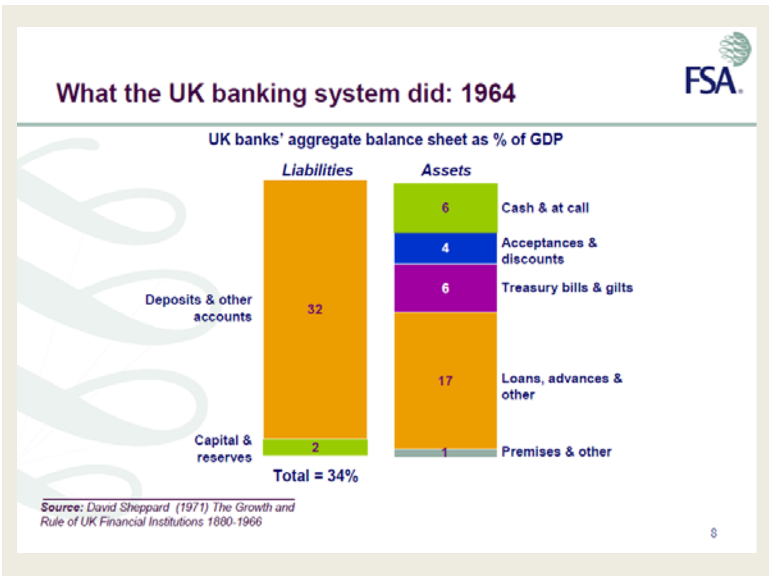
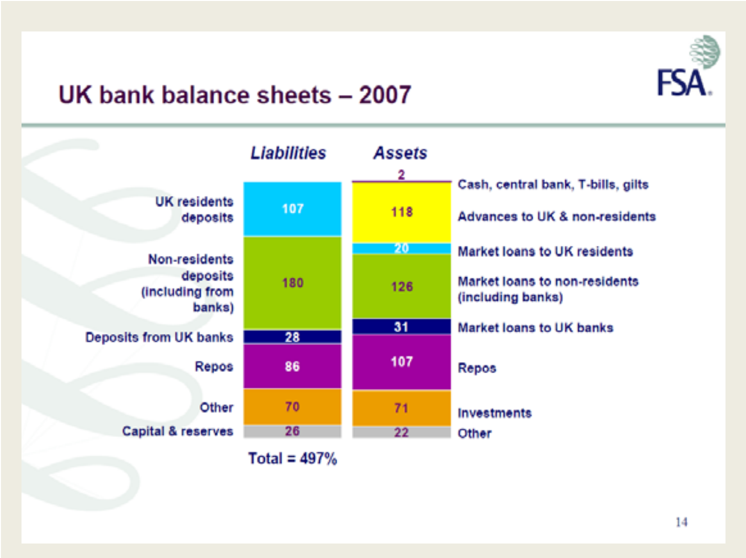


Figure 4

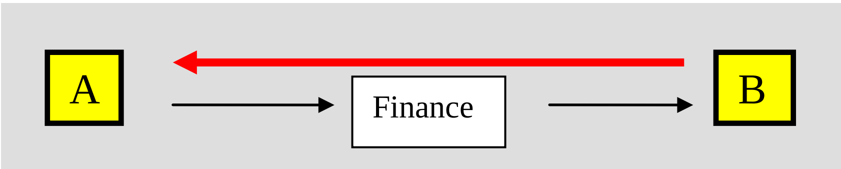


Congestion and the availability of finance

Consider the traditional analysis of congestion, as developed by Pigou (1912). The "underlying" in this case is a journey from A to B. When the road is uncrowded the private cost of a journey is simply the cost in time, fuel and depreciation of the vehicle (ignoring the social cost of some deterioration of the road surface). However, when the road becomes congested the cost must include a systemic component: the congestion cost imposed on every traveller by the presence of others. Note that there is no distinction between the first drivers on the road and the last. All contribute equally to the systemic cost. Moreover the rise in cost is a characteristic of the "system", namely the characteristics of the road from A to B. If the cost of travel from travel directly from A to B becomes excessive, it may well be worth making a detour via C, or perhaps the incentive for diverting to a link via C is the fall in cost associated with a newly constructed highway, replacing the old narrow direct route. The journey is then ACB. The number of links may be multiplied as it becomes more cost efficient to travel via a new diversion rather than direct section CB. So cost minimisation may result in journeys ACDEFB, and so on.

Suppose now that we regard the financing of a transaction as the transportation of the underlying from A to B – say the funding of a mortgage as in Figure 3, (see in schematic form in Figure 5).

Figure 5

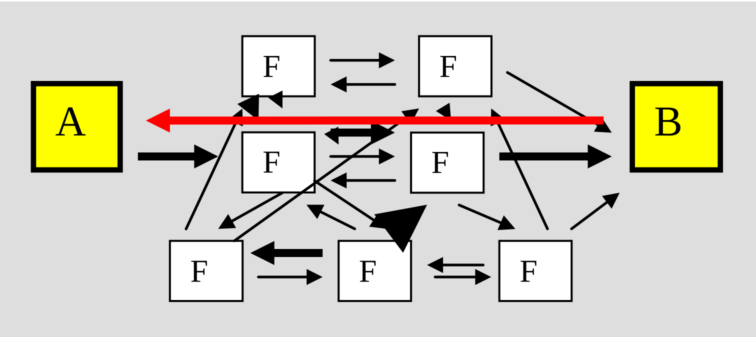


If the value of the underlying is $1, then there are $2 worth of transactions associated with the transfer of the underlying from B to A ($1 A to the bank plus $1 bank to B), and gross asset creation is equal to $2.

The relationship between the availability of funding and the demand for funding is a complex one. In a financial bubble there is typically a severe shortage of liquidity and financial innovation is needed to ease the pressure. The invention of trading on the margin in the coffee houses of Amsterdam in response to the Dutch tulip mania was an early example of such innovation (Posthumus, 1929; Garber, 1989, 1990). This is a direct analogy with traffic congestion. However, causation may well go in the opposite direction. It was the development of credit derivatives that greatly enhanced the incentive to trade tranches of securitised mortgages, now with triple A ratings (Tett, 2009). This was the new super-highway that greatly increased the use of previously underused vehicles (securitisation). There are obvious identification problems as to whether financial innovation fuels spending or a boom in spending stimulates financial innovation. However, the important point for the argument of this paper is that the number, and hence the gross value, of financial transactions associated with a given underlying rises. For the moment it matters not what is the stimulus.

As the demand for funding rises beyond a certain (congestion) point, there is a steady rise in the cost of funding. At a certain point there is an incentive to develop new linkages for funding that will lead to a net reduction in cost (whether through increased liquidity, lower cost or reduced risk). Extra linkages appear in the form of inter-bank transactions and wholesale funding. A schematic illustration is provided in Figure 6. Note that this is more complex than the long chain of Figure 2, attempting to capture the complexity of inter-bank lending displayed, for example, in the repo market.

Figure 6



The growth of short-term interbank linkages increases the number and value of transactions, but the potential increase in marginal cost is reduced, indeed costs may fall.

Define the relationship between financial cost of producing of the underlying of value Y as Y= F(X), and the unit cost as H(x), where x = X/Y is the density of financial services associated with the production of the underling. If h(x) is the cost of applying one extra unit of finance then H(x) = x.h(x). The marginal cost of an increase in the provision of financial services is $H'(x) = h(x) + x.dh/dx$; i.e. the marginal cost is the unit cost plus the change in unit cost due to the increased density of finance x.

If the system of financing stays the same (i.e. as in Figures 1 and 5) then it might be expected that $dh/dx > 0$ as, say, the cost of securing savings from households rises. An extreme example might be a real estate bubble that requires more finance that becomes harder and harder to obtain by traditional means (a modern form of the tulip mania). There is then an incentive to find new ways to produce the required increase in the availability of finance, to reduce its cost, or to reduce its (perceived) micro-risk. New dimensions of wholesale funding are found, and the volume of transactions associated with a given underlying increases. These new transactions (as in Figures 2 and 6) add new links to the chain, equivalent to the new, more complex routings in travelling from A to B, and increase the density of finance. They also comprise, typically, the creation of assets of shorter duration (consider for example the growth of the overnight repo market in the past 20 years).

There is an important dimension to the increase in cost associated with the increase in the density of finance. Consider a simple financial system as portrayed in Figures 1 and 5. Suppose the ultimate borrower defaults on a loan of $100k (and ignore any residual value that may be in the repossessed property that was security for the loan). Then the financial institution has suffered a loss of $100k and cannot meet its liability to the depositor. So, in turn the saving household suffers a loss of $100k, since the household's asset, the bank deposit, is now worthless. In other words the gross asset loss is $200 – the value of the underlying times the number of transactions. Of course it might be argued that the transaction by the financial institution could be netted out and the ultimate loss falls entirely on the saving household. But netting is extremely difficult, because the maturity of the loans is quite different. Yet the whole purpose of intermediation is maturity transformation. So the gross losses dominate the market reaction.

This phenomenon, the impact of gross losses, has been on display in the current financial crisis. For example, Lehman's OTC CDS book had a gross notional of $72 billion and there was initially considerable uncertainty and complexity in identifying close-out positions and replacing defaulted trades. After many months of complex negotiations it has emerged that Lehman's net exposure to OTC CDS contracts was about $5.2 billion. Similarly, in the case of AIG its CDS book had a notional value of $270 billion; whereas actual losses amounted to $3 billion. And its CDOs of ABSs had a notional value of $300 billion and suffered an actual loss of $46 billion. Nonetheless, as asset valuations fell AIG was required to raise additional collateral on the *gross* value of its exposures – something it proved incapable of doing.

So in the event of a systemic failure the potential *gross* loss is a function of financial density: say potential loss is equal to $L(x)$, and $dL/dx > 0$.

Moreover the increase in density increases the *probability* of loss. This is a standard result in the analysis of adaptive routing in loss networks (Kelly, 1991). The analysis of traffic routing (or, similarly, telephone routing) through adaptive networks involves the analysis of the choice of the number of links in the cost minimising journey from A to B (alternatively the routing of a telephone call from A to B):

> "A telephone network provides a fascinating example of a large-scale system where strange effects can occur. For instance, suppose that 'intelligent'

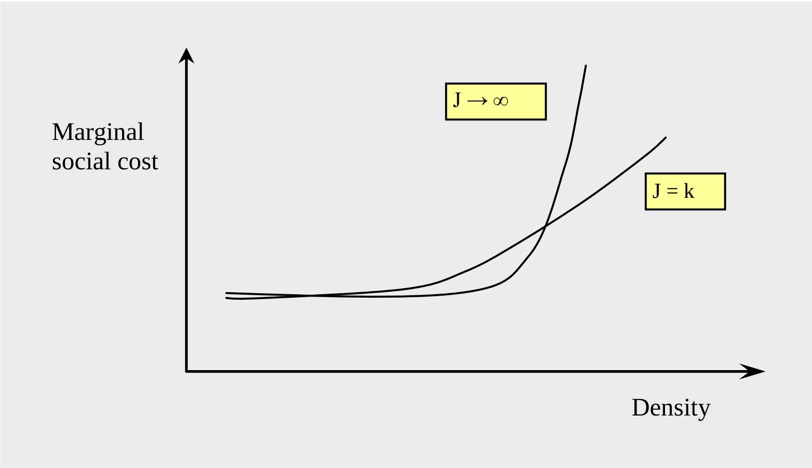
exchanges react to blocked routes by rerouting calls along more resource intensive paths. This in turn may cause later calls to be rerouted, and the cascade effect may lead to a catastrophic change in the network's behaviour. ………. Suppose there is limited intelligence in the form of arithmetical processing ability available for each link *j* and for each route *r*. This intelligence may be located centrally or it may be distributed across the nodes of the network: for example the processing for route *r* might be carried by the source node for calls on route *r*. Suppose that there is limited communication between the intelligences of link *j* and route *r* provided $j \in r$. …..…[Solutions require only that] intelligences know just one item of global information, namely *J*, the total number of links in the system………It is possible to define implied costs and surplus values for fixed point models of alternative routing and trunk reservation…..the potential for long-range disorder and instability is more pronounced in networks with alternative routing, since the chains of influence along different paths tend to reinforce one another. The possibility emphasizes the importance of treating a network as a whole: the local benefits of a capacity or routing change may be completely overwhelmed by adverse consequences elsewhere in the network." (Kelly, 1991, pp. 344 and 362-3).

Translated into the economic case, this becomes:

> "… suppose that 'intelligent' financiers react to expensive funding by developing new means of financing that increases the number of transactions. This in turn may lead to not only to more use of new financial pathways but to further financial innovation, and the cascade effect of a failed transaction may lead to a catastrophic change in the financial system's behaviour. …..the potential for long-range disorder and instability is more pronounced in financial systems with long chains of intermediation, since the chains of influence along different paths of intermediation tend to reinforce one another."

The location of the cost minimising decision (competitive or centralised) is not important, just so long as it is not ordered by an omniscient planner. What is important is the number of links, *J*, generated by the optimising decision. As *J* increases, the probability of catastrophic collapse $P(J)$ rises, $dP(J)/dJ > 0$. Given that the incentive to increase the number of links is a function of x, then $dJ/dx \geq 0$. So the expected value of social loss $S(x) = L(x).P(J)$ and $dS(x)/dx > 0$, i.e. the expectation of social loss (catastrophe) is an increasing function of the financial density (Figure 7).

Figure 7



Summing up: the increase in financial density tends to increase the private cost of disintermediation. This incentivises the search for new financial “routes” or links. But the expected value of catastrophic loss increases with the number of links. This is independent of the ex post structure of linkages. It is purely “systemic”.

Policy

The social cost of an increase in financial density is a sum of the increase in private cost and the expectation of gross loss (given that the possibilities of netting are very limited). Private marginal cost may well be constant or falling, creating the socially perverse incentive to increase financial density, which increases the probability of catastrophe. It is a standard result (Pigou, 1920) that externalities, such as the expected cost of catastrophic collapse, should be recognised by a charge on the density of “traffic”, in this case on the density of financial transactions relative to the value of the underlying. The charge should be “anonymous”, i.e. there is no reason to discriminate between particular financial transactions, all are culpable.

Hence the case for a financial transactions tax may be made entirely independently of any considerations of “short-termism” or ex post characterisations of the pattern of linkages within the financial system. It may be based solely on the level of financial density, as manifest in the number of linkages associated with the financing of any given underlying (such as gdp).

There is also a strong case for an increase in the possibility of netting, resulting in a reduction in L(x). The maturity transformation that lies at the heart of banking is particular acute in relation to short-term wholesale funding of longer term commitments (assets). Whilst in principle deposits by households are demand deposits, they are typically very “sticky”; hence their characterisation as “core” liabilities (Shin, 2008 and 2010). Wholesale funding that relies on continuous market re-financing is far less sticky. Indeed the persistent emphasis on the enhancement of liquidity in the financing process increasing the potential volatility of funding. A way to manage this volatility is to more nearly match the maturity structure of liabilities and assets. This also increases the possibility of netting. Recent regulatory proposals characterised as increasing liquidity, are in fact designed to match the maturity of liabilities to the maturity of assets. The proposed Liquidity Coverage Ratio requires that the ratio of the stock of high quality liquid assets to potential 30 day net cash outflows should exceed 100%. The proposed Net Stable Funding Ratio requires that longer term assets are matched by customer deposits, *long-term* wholesale funding and equity. Any reduction in maturity transformation is likely to reduce the ability of the banks to make money. It will also reduce L(x).

References

Adrian, T. and M. Brunnermeier, (2009) *CoVaR*, mimeo, Princeton University.

Allen, F., A. Babus and E. Carletti (2010) "*Financial Connections and Systemic Risk*", Working Paper 10-20, Wharton Financial Institutions Center, University of Pennsylvania.

Allen F. and A. Babus (2009) "Networks in Finance", in *The Network Challenge*, edited by P. Kleindorfer and J. Wind, Wharton School Publishing, University of Pennsylvania, pp. 367-382.

Brunnermeier, M., A. Crockett, C. Goodhart, A. Persaud and H.S. Shin. (2009). *The Fundamental Principles of Financial Regulation,* CEPR, London.

Garber, P.M. (1989) "Tulipmania", *Journal of Political Economy.*

Garber, P.M. (1990) "Famous first bubbles", *Journal of Economic Perspectives.*

Haldane, A. (2009) *Rethinking the Financial Network*, Bank of England, London.

Kelly, F. (1991) "Network routing", *Royal Society Philosophical Transactions: Physical Sciences and Engineering*, Vol. 337, No. 1647, Complex Stochastic Systems, pp. 343-367.

Pigou, A.C. (1912) *Wealth and Welfare,* Macmillan, London.

Pigou, A.C. (1920) *The Economics of Welfare,* Macmillan, London.

Posthumus, N.W. (1929) "The tulip mania in Holland in the years 1636-37", *Journal of Economic and Business History.*

Shin, H.S. (2008) *Risk and Liquidity*, 2008 Clarendon Lectures in Finance, Oxford University Press forthcoming

Shin, H.S. (2010) *Macroprudential Policies Beyond Basel III*, mimeo, Princeton.

Tett, G. (2009) *Fool's Gold: How Unrestrained Greed Corrupted a Dream, Shattered Global Markets and Unleashed a Catastrophe,* Little, Brown, London.

Vickrey, W. (1969) "Congestion theory and transport investment", *American Economic Review.*

Walters, A.A. (1961) "The theory of measurement of private and social costs of highway congestion", *Econometrica.*