1. Summary

This paper studies the welfare implications of various government policies on the banking industry: suspension of convertibility, taxation on short term deposits, reserve requirements, partial or total deposit insurance and capital requirements. Bank stability has always been a major concern for regulators, policymakers and among academics, as the banking system has traditionally been vulnerable to the problem of banking panics. Examples of financial crises in the history of the financial systems were the Great Depression (1929–1933) which had a significant impact on the banking system of the US1 or the more recent crises in emerging countries.2 Given the historical importance of banking panics and their current relevance, it is important to understand why they occur and what policies should be implemented to deal with them. In this sense, the theoretical research on banking has focused on analyzing the microeconomic nature of banks and their role in the economy. Diamond and Dybvig (1983), which formalized some of the ideas provided in Bryant (1980), made a significant contribution by modeling the demand for liquidity and the transformation service provided by banks.

The authors consider a simple model with a continuum of ex ante identical agents who are risk averse and uncertain about their future time preferences. These individuals are born with one unit of the good at $T = 0$ and are subject at $T = 1$ to privately observed risk.

---

1. From 1930 to 1933 the number of bank failures in the US averaged over 2000 per year, see Mishkin (1995).
2. Lindgren et al. (1996) find that 75 per cent of the IMF’s member countries suffered banking crises between 1980 and 1996.
with probability $\alpha$ of being type 1 who derive utility only from consumption in the first period and probability $1 - \alpha$ of being type 2 who derive utility only from consumption in the second period. There is an investment technology such that a unit investment at $T = 0$ yields one unit at $T = 1$ or $R > 1$ units at $T = 2$. In autarky, early consumers liquidate their investment at $T = 1$ and consume one unit whereas late consumers maintain the technology and receive $R > 1$ at $T = 2$.

Diamond and Dybvig show how a financial intermediary can improve ex ante welfare by offering consumers a deposit contract. Suppose that banks operate in a competitive environment; hence, the optimal contract is the one that maximizes the expected utility of the agents subject to a zero profit constraint. This deposit contract promises investors a higher payoff for early consumption and a lower payoff for late consumption compared to autarky. Type 1 agents are insured against being early consumers in the sense that they receive some of the benefits available from the long term technology.

In the original Diamond and Dybvig model, there are two Pareto-ordered Nash equilibria: a Pareto dominant equilibrium that achieves socially optimal risk sharing which has only type 1 agents withdrawing at $T = 1$, and a second Pareto dominated equilibrium in which all agents withdraw at $T = 1$ and can be interpreted as a bank run. Finally, the model shows that there are several measures to prevent the bank run equilibrium.\(^3\)

However, this seminal paper attracted an important criticism for assuming that bank runs are random phenomena, and thus, uncorrelated with other economic variables. Gorton (1988), in an empirical study of bank runs in the US during the National Banking Era (1863–1913), found support for the notion that bank runs tended to occur after business cycle peaks.

Second, models of information-induced runs assert that bank runs occur due to the diffusion of negative information among depositors about bank’s solvency, that is, bank runs are related to the state of the business cycle, as supported by the empirical evidence. Examples of this literature are Chari and Jagannathan (1988), Jacklin and Bhattacharya (1988), Alonso (1996) and Allen and Gale (1998), among others.\(^4\)

Zhu’s paper follows the business cycle view, as to the origins of banking panics. The model is based on Allen and Gale (1998). There are three periods. Two assets are available in period 0: a storage technology and a risky one. The riskless asset transforms one unit of consumption in any period $t$ to one unit of consumption in a period $t + 1$. The risky asset yields a random return $\tilde{R}$ in period 2, and can be liquidated prematurely at a cost. On the consumer side, there is a continuum of depositors, that are born with one unit of endowment in year 0, and are subject to a preference shock in year 1. As in Diamond and Dybvig, they can be of type 1 with probability $\alpha$ and of type 2 with probability $1 - \alpha$.

As there is no aggregate uncertainty in the model, a fraction $\alpha$ of agents will be of type

---

\(^3\) In the case in which there is no aggregate uncertainty, a suspension of convertibility policy in which withdrawals up to $\alpha$ are allowed would implement the good Nash equilibrium. This policy removes the incentives of type 2 depositors to withdraw early, as independently on what other agents do, they always obtain a higher payoff if they wait until the second period than if they withdraw. However, if there is aggregate uncertainty, this measure may not be so effective, depending on the realized $\alpha$. In this case, they advocate for Federal Deposit Insurance (guaranteed by governmental funds) as the effective mechanism that would implement the Pareto dominant equilibrium.

\(^4\) It should be mentioned that there are a number of recent papers that study banking panics in the context of the entire banking system, as Chen (1999), Freixas et al (2000), Allen and Gale (2000) or Aghion et al (2000).
1 and a fraction $1 - \alpha$ of type 2. Finally, there is a perfectly competitive banking system that offers individuals a demand deposit contract, in exchange for their endowment. This contract provides a risk sharing mechanism against uncertainty in liquidity needs. It will specify both a short term interest rate ($r_1$) and a long term one ($r_2$). Formally, $r_1$ and $r_2$ are obtained by maximizing the expected utility of agents subject to resource and incentive compatibility constraints. It should be noticed that the long term interest rate will be state contingent (as the bank invests in a random technology, it may not be able to guarantee the promised long term interest rate), and so there is a threshold value of the random return, $R^*$, below which the contract is no longer incentive compatible and therefore type 2 agents will prefer to withdraw in the first period. In fact, when the return is low, all type 2 agents will run on the bank and will force the bank to liquidate its assets at a loss. Therefore, bank runs are costly and destroy the risk sharing mechanism.

The allocation achieved in this benchmark model is compared with two types of socially optimal allocations: the first one is the Allen and Gale (AG) optimum, which diversifies the idiosyncratic preference shock and involves no liquidation of the risky asset. It is obtained by maximizing the expected utility of agents subject to resource and incentive compatibility constraints:

$$\max_{c_1(\mathcal{R}), c_2(\mathcal{R}), x} \left[ E_0 \left[ \alpha U(c_1(\mathcal{R})) + (1 - \alpha) U(c_2(\mathcal{R})) \right] \right]$$

subject to:

$$\alpha c_1(\mathcal{R}) \leq 1 - x$$
$$c_1(\mathcal{R}) \leq c_2(\mathcal{R})$$
$$1 - \alpha c_2(\mathcal{R}) + \alpha c_1(\mathcal{R}) \leq xR + 1 - x$$

where $c_1(\mathcal{R})$, $c_2(\mathcal{R})$ represent consumption in years 1 and 2 respectively and $x$ is the investment in the risky asset. There are two possible solutions: When the return is low, the optimal allocation involves carrying over some of the liquid asset to year 2, to supplement the low realization of the risky asset. All agents receive an equal payment.\(^5\)

$$c_1(\mathcal{R}) = c_2(\mathcal{R}) = xR + 1 - x$$

On the contrary, when the return is high, the bank exhausts the liquid asset in year 1, as consumption in year 2 will be high, in any case.\(^6\)

$$c_1(\mathcal{R}) = \frac{1 - x}{\alpha}$$
$$c_2(\mathcal{R}) = \frac{xR}{1 - \alpha}$$

It should be noticed that in this allocation, the aggregate risk related to the risky asset still exists, as the consumption levels are state contingent. However, if the social planner could

---

\(^5\) The condition for this case to hold is obtained by substituting the optimal consumption levels in the first resource constraint, that is: if $\alpha c_1(\mathcal{R}) \leq 1 - x$ or equivalently, $R \leq ((1 - x)(1 - \alpha)) / (\alpha x) = R^*$

\(^6\) Similarly, the condition for this case to hold is obtained by substituting the optimal consumption levels in the incentive compatibility constraint, that is: if $c_1(\mathcal{R}) \leq c_2(\mathcal{R})$ or equivalently, $((1 - x)(1 - \alpha)) / (\alpha x) = R^* \leq R$
use the profits in good states to subsidize the interest payment in bad states, the aggregate risk would be removed, and agents would receive a fixed consumption regardless of the state of the economy. This allocation would be the first best optimum, and is characterized by:

\[
\begin{align*}
  c_1 &= \frac{1 - x}{\alpha} \\
  c_2 &= \frac{x E(R)}{1 - \alpha}
\end{align*}
\] (5)

However, in this case, one has to assume that the social planner has outside resources, not available to private agents, so that he deliver the expected value.

Once the socially optimal allocations have been defined, the paper analyzes different regulation measures to prevent banking panics. The above mentioned framework enables the author to consider the ex ante and ex post effect of the different instruments: Suspension of convertibility, reserve requirements and taxation on short term deposits turn out to be both ex ante and ex post inefficient to prevent banking panics. They produce either a misallocation effect in the case of suspension or distortions in the bank’s investments decisions, in the case of the other two policies. On the other hand, a full deposit insurance system would be ex post efficient to prevent banking panics, but is ex ante inefficient due to the moral hazard problem. In this sense, an interest-cap deposit insurance scheme overcomes the moral hazard problem and achieves the first best optimum. Finally, capital requirements may also achieve the first best optimum, provided the level of capital is high enough.

2. Comments

The paper deals with an interesting and important topic, which is analyzing the causes of banking panics so as to make adequate policy prescriptions and does a complete analysis of the main existing regulatory policies. There are few theoretical papers\(^7\) that have analyzed the ex ante effects of banking regulation, and in this sense, this paper is an interesting contribution to the theoretical literature.

As already mentioned, this paper is based on Allen and Gale (1998)(hereafter AG). However, the main difference with respect to the AG model, is that in this paper, the risky asset can be liquidated prematurely at a cost. This assumption is crucial in the paper for the results obtained herein:

The illiquidity assumption in AG (first section of the paper) plays an important role in that model: first, bank runs are always partial. As long as the value of the risky asset is greater than zero there is always a fraction of type 2 consumers that wait to withdraw in year 2, as otherwise second period consumption would become infinite. Second, demand deposit contracts and bank runs achieve the first best allocation (AG optimum). In other words, when bank runs occur in equilibrium, there is always a unique proportion of type 2 consumers (less than the total one), given by: \(\alpha_1(R) = ((1 - x)/(1 - x + xR)) - \alpha\), that equates the consumption of both types of agents and therefore replicates the optimal allocation. Finally, pure panic runs are always excluded.

\(^7\) Recent examples are Freixas and Gabillon (1999) or Samartin (2002)
On the contrary, when the risky asset can be liquidated prematurely at a cost, bank runs are no longer partial, they involve all type 2 consumers and cause the costly liquidation of assets. As a result, demand deposits and bank runs do not longer achieve the AG optimum. Finally, pure panic runs need to be eliminated using an equilibrium selection mechanism.

A second point is related to the specification of the AG optimum, in this paper. Why is not liquidation specified as a possibility although it is not optimal? The author assumes an exogenous and very low liquidation value of the risky asset, and therefore, it is true that liquidation turns out not to be optimal. However, partial liquidation of the risky asset may be optimal for greater liquidation values. In this respect, the author might look at a paper by Hellwig (1994), where he finds that premature liquidation of the long-term technology is optimal for the second best allocation. And a related point is that as a possible extension one could assume an endogenous liquidation value (the last section of the Allen and Gale paper may also shed some light on this direction).

Regarding the welfare analysis of the different policies, and in particular the suspension policy, it is an intervention measure that was frequently used in the U.S. during the nineteenth and early twentieth centuries and consists on the banks refusal to exchange currency for demand deposits upon demand. Gorton (1985) points out that during this period convertibility was suspended up to eight times in the US.

An interesting aspect is that, even though this measure implied a violation of the deposit contract, neither depositors nor the banks were against it. It was understood as a temporal measure to stop banking panics. During this period, the bank tried to solve the liquidity problem and with the suspension, depositors recovered the confidence in the system by recognizing the institutions in trouble. Hence, by suspending convertibility, banks can signal to depositors that continuation of investments is mutually beneficial.

Another important era in which suspension worked quite well, was the Scottish free banking era. Scotland’s free banks were incredibly disciplined as a result of suspension, spurring England to pass protective legislation to keep Scottish banks out of the adjoining market.

The author argues that this measure is not ex ante efficient because it causes a misallocation effect as there is no guarantee that only those agents that have true liquidity needs will get their payment in year 1. It should be mentioned, however, that in this model, no sequential service constraint is assumed, and so one could consider that funds are distributed equally among agents that withdraw in the first period. And also, why do banks suspend convertibility at a level of $\alpha$? In this sense, the author might consider what could be called a state contingent suspension policy. Under this measure, withdrawals up to a level $\alpha + \alpha_1(R)$ are allowed, where $\alpha_1(R) = ((1 - x)/(1 - x + xR)) - \alpha$, represents the fraction of type 2 agents that withdraw in the first year, and that depends on the realization of $R$. In other words, this suspension is equivalent to allowing for partial bank runs, and therefore it would achieve the AG optimum. Therefore, proposition 2, is misleading, because suspension can be ex ante efficient, if applied at the right level.

With respect to taxation on short term deposits and liquidity requirements, it should be mentioned that the inefficiencies of both policies are also stated in AG paper.

The author also considers a full coverage deposit insurance system, or blanket guarantee scheme. The bank pays a premium to the insurance fund, and has its deposits insured. In this framework, the bank promises depositors interest rates that are no longer state
contingent, and when its return turns out to be low, depositors are paid by the insurance fund. It is assumed that the bailout costs can be covered by the insurance premium payment, or equivalently, the insurance fund or public authority always has enough funds. Therefore, the social cost of deposit insurance is zero, or equivalently, the depositor supports no deadweight taxes. Perhaps, this point should be made clearer, as it is key factor in welfare analysis.

Also, as it was pointed out above, the author assumes that the insurance provider can cover actual losses by providing resources. If the model is interpreted as a three-period general equilibrium model, this is not possible. Having an insurance premium that equals the expected payout does not work unless there are more periods or more resources. This is exactly the same as the specification of the first best allocation. One has to assume that the government can introduce outside resources.

An interesting result is that partial deposit insurance removes the well known moral hazard problem, associated to full deposit insurance, and achieves the first best allocation. Recent empirical studies have also pointed out that explicit insurance systems create more stable banking systems, see for example Gropp and Vesala (2001), and in reality many countries do have partial deposit insurance. The problem is as usual the too-big-too-fail policy, associated to big banks, that are sure of having an implicit coverage. Although this issue cannot be analyzed in this framework with a perfectly competitive banking system, it might represent an interesting extension for future research. Another question is whether it could be possible to construct a risk based insurance premium.

Finally, the paper looks at another widely used tool in bank regulation, which is the imposition of a capital requirement. In the model, there are other risk-neutral investors who provide outside resources, to fulfill the capital level. This policy converges to the first best optimum as the capital ratio approaches the 100 per cent ratio. This implies very high levels of capital, as the author mentions, and there is also the potential problem of a credit crunch, as several authors have documented. Similarly to deposit insurance, capital requirements also assume the introduction of outside resources. One way to avoid these problems, would be to assume an infinite-horizon setting, either with infinitely-lived agents or overlapping generations. Alternatively, one could use explicit partial-equilibrium welfare measures based on consumer’s surplus.

Obviously, all these extensions are beyond the scope of one unique paper, but nevertheless, represent interesting lines for future research. In this respect, Zhu’s analysis points the way to possible resolutions of important issues in banking regulation.

Acknowledgements

The author thanks Gerald Dwyer, organizer of the Regulation and Financial Stability Conference, held in Atlanta (2004), as well as the editors of the Journal of Financial Stability for their invitation to present and publish this comment.

References

Allen, F, Gale, D, 1998 Optimal financial crises J Fin 53, 1245–1284
Allen, F, Gale, D, 2000 Financial contagion J Pol Econ 108 (1), 1–33
Alonso, I, 1996 On avoiding bank runs J Monet Econ 37, 73–87
Bryant, J, 1980 A model of reserves, bank runs and deposit insurance J Bank Fin 4, 335–344
Chari, V, Jagannathan, R, 1988 Banking panics, information and rational expectations equilibrium J Fin 43 (3), 749–761
Chen, Y, 1999 Banking panics: the first come, first served rule and information externalities J Pol Econ 107, 946–968
Diamond, D, Dybvig, P, 1983 Bank runs, deposit insurance and liquidity J Pol Econ 91 (3), 401–419
Freixas, X, Gabillon, E, 1999 Optimal regulation of a fully insured deposit banking system J Regulat Econ 16, 111–134
Freixas, X, Parigi, B, Rochet, C, 2000 Systemic risk, interbank relations and liquidity provision by the central bank J Money, Credit Bank 32 (3), 611–638
Hellwig, M, 1994 Liquidity provision, banking and the allocation of interest rate risk Eur Econ Rev 38, 1363–1389
Lindgren, C, Garcia, G, Saal, M, 1996 Bank soundness and macroeconomic policy International Monetary Fund
Samartin, M, 2002 Suspension of convertibility versus deposit insurance: a welfare comparison Eur Fin Rev 6 (2), 223–244