Principles for macroprudential regulation

Dimitrios P. Tsomocos
University Reader in Financial Economics, Said Business School
Fellow in Management, St Edmund Hall, University of Oxford

Alexandros Vardoulakis
Economist
Board of Governors of the Federal Reserve System

The drafting of macroprudential regulation is largely being driven by the need by policy makers to meet timetables that have been agreed. The legislative drive is taking place without any clear theoretical framework to organise the objectives. In this article we propose two principles that any satisfactory framework ought to respect and then describe one specific model that embodies these principles. We explain the insights from this approach for regulatory design.
Under our current system of safety-and-soundness regulation, supervisors often focus on the financial conditions of individual institutions in isolation. An alternative approach, which has been called systemwide or macroprudential oversight, would broaden the mandate of regulators and supervisors to encompass consideration of potential systemic risks and weaknesses as well.” Ben Bernanke, August 22, 2008.

1| GUIDING PRINCIPLES FOR MACROPRUDENTIAL REGULATION

The quote above from Ben Bernanke, about three weeks before the failure of Lehman Brothers delivered at Jackson Hole in front of the top central bankers around world, shows regulators even before the intense phase of the global financial crisis (GFC) recognised that the existing regulatory architecture was deficient. In the subsequent five years, while many speeches, working groups and proposals have been made to remedy the problems, tangible progress has been relatively modest. Bank regulations involving capital requirements have been substantially refined. Other changes to rules regarding liquidity, or to give regulators other powers over, say dividends or loan-to-value (LTV) standards have been delayed. Indeed, it is fair to say that there is still no unified framework that organises thinking about how to proceed.

In this article we step back from the immediate policy debate to propose two principles about how to organise future discussions of macroprudential regulation. We derive these principles starting from a pair of definitions proposed by Eric Rosengren (2011) that we believe captures well the sense of many experts:

“Financial stability reflects the ability of the financial system to consistently supply the credit intermediation and payment services that are needed in the real economy if it is to continue on its growth path.”

“Financial instability occurs when problems (or concerns about potential problems) within institutions, markets, payments systems, or the financial system in general significantly impair the supply of credit intermediation services – so as to substantially impact the expected path of real economic activity.”

We like Rosengren’s definitions for three reasons. First, they rightly place the goal of delivering financial stability as supporting the real economy. This immediately implies that activities that are superfluous to supporting growth are rightly seen as non-essential ones that could be jettisoned. Second, he takes a broad view of what the financial system does to support the economy. As we explain below his focus on intermediation and payment services matches well with the academic literature on this issue. Third, he emphasises threats that come from not only problems that actually materialise, but also those that simply might arise and emphasises the need to guard against both. We will see that this prescription also fits well with an analytic approach to regulatory design.

To operationalise Rosengren’s definition, it is necessary to determine how offering intermediation and payment services supports economic activity. Unless we are precise about these channels then it is impossible to judge whether regulations which will restrict the system are on balance worthwhile. A corollary to this perspective is that being precise about the reasons why regulation is needed in the first place is also useful for guiding regulatory design.

There is a vast academic literature on the social purpose of the financial system which points to three contributions. The first is to expand the amount of credit that can be extended to a given borrower. The micro-founded explanations for this conclusion typically assume that borrowers can potentially default on loans and so any lender has to be diligent in monitoring borrowers (Diamond, 1984). By concentrating the lending with specialised agents these monitoring costs can be conserved and the amount of credit extended can be expanded.

A second widely posited role for the financial system is helping people and businesses share risks (Benston and Smith, 1976; and Allen and Gale, 1997). There are many ways to formalise how this takes place, but one simple one is to recognise that by having banks that not only offer deposits, but also allow savers to buy bank equity, the banks can create two different types of claims that would be backed by risky loans. These two choices allow savers to hedge some risks associated with lending and this hedging improves the consumption opportunities for savers.
A third role, which the literature recognises as complimenting the second, is having a financial system that creates liquid claims that facilitate transactions. There are various motivations for how this can be modeled. In Diamond and Dybvig (1983), a financial intermediary can cross-insure consumers’ needs for liquidity by exploiting the law of large numbers among customers. But doing so exposes banks to the possibility of a run which can be disastrous for the bank and its borrowers and depositors. Calomiris and Kahn (1991) and Diamond and Rajan (2001) explain that the very destructive nature of the run is perhaps helpful in disciplining the bank to work hard to honor its claims. So the fragility associated with the runs is potentially important in allowing both high amounts of lending and large amounts of liquidity creation.

These observations bring us to our first principle:

**Principle No. 1.** Any satisfactory framework for analysing financial stability and macroprudential regulation should be rich enough to account for all three of these contributions of the financial system.

It is perhaps easiest to appreciate the importance of this principle by looking at some examples which violate it. One prominent policy proposal that is often held up as an appealing alternative to current bank regulations is to insist that banks hold only liquid securities as assets (Kotlikoff, 2010). The creation of these narrow assets would eliminate the risk of a bank run and still allow banks to provide liquid assets to their customers. One might imagine that banks could even be profitable if they were buy large denomination securities and using the law of large numbers to manage the transactions costs of selling them when cash is needed for customers.

This kind of a proposal is superficially appealing but it excludes the credit creation function of the banking system. Hence, if implemented, it would force that intermediary can cross-insure consumers’ needs for liquidity by exploiting the law of large numbers among customers. But doing so exposes banks to the possibility of a run which can be disastrous for the bank and its borrowers and depositors. Calomiris and Kahn (1991) and Diamond and Rajan (2001) explain that the very destructive nature of the run is perhaps helpful in disciplining the bank to work hard to honor its claims. So the fragility associated with the runs is potentially important in allowing both high amounts of lending and large amounts of liquidity creation.

Likewise, in the aftermath of the GFC many economists have argued that bank capital requirements are far too low. An extreme example of this is Admati, DeMarzo, Hellwig and Pfleiderer (2010) who advocate creating a banking system that creates no liquid claims against its risky assets. Again this bank will be free from runs and able to absorb credit losses without needing any taxpayer support. But, the arrangement supposes that liquidity provision is not a core function of banks and that precluding them from providing liquidity is costless. So this type of analysis also strikes us as incomplete and ill-suited as a starting point for regulatory design.

In what follows we will sketch a model developed by Kashyap, Tsomocos and Vardoulakis (2014), henceforth KTV, that includes all these three roles mentioned above for the financial system. In the KTV model, there are three underlying frictions that influence that way transactions are structured and the social contribution of the financial system. One is the incompleteness of contracts, i.e. not all potential contingencies can be spelled out in advance. This means that are some circumstances where agents may default inadvertently (or perhaps strategically) and potentially these defaults spillover to affect others parties in the economy.

A second friction in the KTV model is the assumption that borrowers are subject to limited liability. A popular narrative about the GFC holds that some large institutions took reckless gambles knowing that they would not be held fully accountable for any associated losses. By building in limited liability for contracts directly into the model this incentive is present.

The third friction is borrowers do not fully appreciate the consequences of their actions on the interest rates that they face. This is a much more subtle factor than the other two and we explain it in detail below. But, it is important because it determines the equilibrium cost of credit in the economy. One of the strengths of the Rosengren approach is to recognise that potential threats to stability, not just realised ones can matter. To fully account for both types of risks, forward looking behaviour must be present and prices must reflect these future possibilities.

This brings us to our second principle:

**Principle No. 2.** Any satisfactory framework for analysing financial stability and macroprudential regulation should incorporate forward looking behaviour and have prices that adjust endogenously to reflect potential risks.
At a technical level, this principle implies that the analysis should be conducted in a general equilibrium environment so that the endogenous actions that agents take to counteract the various distortions are reflected in prices. There are several other reasons to favor a general equilibrium approach. By deriving behaviour from utility maximisation with rational expectations and market clearing, we ensure that the model will specify coherent behaviour, even if policy essentially changes the regime in which an agent operates. General equilibrium also makes welfare analysis, and especially distributional questions, amendable to a rigorous treatment. In general equilibrium one sees all the indirect effects and feedback mechanisms, so that one can judge whether, and under what conditions, they can be safely ignored. So although the effects we describe below depend on the exact parameterisation of the model, the ones we emphasise are relatively robust and their strength varies quantitatively, but not qualitatively.

At a practical level, this principle opens up a role for both ex ante regulation that operates on incentives to avoid problems and ex post policies that seek to mitigate the damage after bad realisations occur. Many policy discussions can be confusing because the environment is not rich enough to allow for both these types of policies. By respecting this principle we can be assured that the framework does not presuppose that either the ex ante or ex post approach is necessarily better.

We now turn to a description of the KTV model of financial intermediation that embodies these two principles. We view this model of more of an example of the benefits from adopting this particular approach to thinking about macroprudential regulation, rather a fully calibrated model that would be suitable for using in a quantitative analysis of competing regulations. Hence, in explaining the model we emphasise the intuition behind the way it operates and its qualitative predictions. That treatment is sufficient to allow for a high-level comparison of competing regulations even in the absence of a tightly parameterised, empirical mode.

2| **A SPECIFIC MODEL OF FINANCIAL INTERMEDIATION**

The framework proposed by KTV is an extension of the widely studied three-period model that was first developed by Diamond and Dybvig (1983) (and explained with a very accessible and intuitive example in Diamond, 2007). In Diamond and Dybvig (1983) savers have an uncertain demand for access to liquid assets and the economy naturally delivers opportunities to invest in a highly illiquid asset. The loan is not risky in that it has a guaranteed high rate of return if the project is allowed to continue to completion; so a loan initiated in period 1 can pay back with certainty in period 3. If the lender asks for early repayment of the loan in period 2, however, the amount that can be recovered is far below the promised level. The problem is that some depositors might have an urgent need for funds in period 2. Consequently, if any individual saver opts to make a direct loan to fund an illiquid project and then discovers that funds are needed in the interim period, the early repayment will be very low. So by assumption direct lending leaves savers with very little available resources in the event that liquidity is needed on short notice.

Banks in their model pool the risks regarding liquidity needs over many customers. So the bank has no advantage in forecasting liquidity needs, or even in the proceeds it gets from demanding early repayment on any particular loan. But, the bank can guess that in normal circumstances most savers will not need their money on short notice. So the bank can make the following offer to savers: compared to directly lending yourself, we offer a deposit contract that pays you more in the event that you want to get your deposit back in period 2 and less than if you can get if you wait until period 3 to be repaid.

The whole reason for impatience in normal circumstances is an urgent need for funds, so the marginal utility of having the money in the second period is high on occasion. Diamond and Dybvig assume every depositor has some chance of needing money in the interim period and no one knows as of period 1 whether they will need the money early or not. Accordingly, reducing the variance of the return on deposits by giving up some of the extra return from being patient in exchange for a higher return when a depositor is impatient is appealing to the depositor.

The bank can support this contract by calling in extra loans to pay off the impatient depositors, so for each depositor who asks for withdrawal in period 2, the bank recalls more than one loan to make that repayment. With fewer loans outstanding, the total amount available in period 3 will be less than the
number of depositors who turn out to be patient. Diamond and Dybvig suppose that the number of depositors (and borrowers) is large, so that the bank can make an informed guess about the aggregate number of impatient depositors and set the returns promised to the patient depositors low enough to credibly pay them and the impatient depositors. Hence, the bank is in the business of offering liquidity insurance.

The problem with this arrangement is that if an unexpectedly large number of depositors request an early withdrawal, then the arrangement can unravel. At a certain point every patient depositor can realise that too much has been paid out to people withdrawing early, so that by period 3 the full promised return will not be available. If all the patient deposits could coordinate, they would agree that waiting and taking whatever is available gives them more than running and seeking an early withdrawal. When they all line up to get their deposits back, some will wind up getting paid the lower promised return for an early withdrawal and others get nothing when the bank has called in all its loans. Unfortunately, being patient is not individually rational if a depositor conjures that everyone else is going to run. So the Diamond-Dybvig bank is inherently fragile and vulnerable to a run that leads to some depositors not being paid and a collapse in credit extension where loans are called before projects come to fruition.

KTV modify the basic Diamond and Dybvig setup in five important ways so that the extended model satisfies the two principles proposed in section 1. First, they introduce three types of agents: savers, bankers, and entrepreneurs. Each of these agents is given an endowment in period 1 that can be consumed or invested. In addition, the banker begins with some equity that is trapped in the bank and can only be invested. This creates a natural reason for the banker to take funds from the saver and function as an intermediary.

A second change is that the entrepreneurs have unique access to a risky investment opportunity. The entrepreneurs have insufficient funds to operate at the efficient scale so they must borrow. The banks have a natural advantage at lending and can lend more than would be available if the entrepreneurs had to borrow directly from the savers. The projects that are financed have a stochastic payoff, so unlike in Diamond and Dybvig there is intrinsic investment risk in the economy.

Third, KTV posit that banks and entrepreneurs are subject to limited liability. This creates an incentive for both the banks and entrepreneurs to take excessive risk, which translates to undertaking more risky investment than otherwise. So it is possible that a bad outcome will arise because of this gambling to exploit the limited liability.

KTV’s fourth change is that savers face a portfolio decision in which they can directly invest in a safe asset or invest in the bank in the form of either deposits or equity. In Diamond and Dybvig the safe asset is dominated by making loans and if necessary liquidating them early. In KTV, the saver could choose to hold the safe asset to avoid the excessive gambling by the banks and entrepreneurs. The banks can also invest in the safe asset if the discount from calling in the loan early is sufficiently high.

Finally, KTV make an assumption about how depositors decide whether or not to run. In the Diamond and Dybvig model, a run can occur because of a pure failure to coordinate by the patient depositors; in the jargon that is now popular, a random event like a “sunspot” could lead to a panic where all the patient depositors decide to run (Cass and Shell, 1983). KTV instead make the probability of a run a random variable which is more likely to occur based on the fundamental condition of the bank. When the bank has more risk because of increased lending, the probability of a run increases. Likewise, when the bank is more levered, i.e. has higher deposits relative to equity, a run is also more likely.

Together these ingredients are enough to create a very rich environment that facilitates both a role for regulation and different outcomes depending on the form that regulation takes. One important consideration is that there are still not enough assets for savers to fully hedge the risks that they face.

---

1 Goldstein and Pauzner (2005) show that this kind of result can be derived formally in a model that is in the spirit of Diamond and Dybvig (1983).

In the Goldstein and Pauzner setup the savers each receive noisy signals about the health of the bank and must form beliefs about whether the other depositors will run. They show that an equilibrium of the sort constructed by Morris and Shin (1998) exists, i.e. depositors will run only when their signal is below a threshold determined by fundamentals. Thus, the ex ante probability of a bank-run is computed uniquely as a function of other endogenous variables.
So they need to choose how much to save using deposits, bank equity and the safe asset to control the risk taking incentives created by limited liability.

The distorted incentives of the banks and entrepreneurs to gamble also affect outcomes in a second more subtle way. The limited liability puts a floor on private payoffs to the bank and the entrepreneur from taking more risk. When the saver makes an investment in the bank, the saver understands the risk taking incentives of the bank and demands appropriate risk adjusted returns to compensate for the risk. Likewise, when the bank lends to the entrepreneur, the bank recognises the entrepreneur's temptation to adjust risk and appropriately prices its loan to account for this.

So although the interest rates (and required return on bank equity) reflect the risk in the economy, both the bank and the entrepreneur take the prices as given. Hence, neither the bank nor the entrepreneur fully internalises the effect that their own actions have on the prices they face. As is usually the case in these situations, this means that the privately optimal choices made by the bank and the entrepreneur will not match the socially optimal choices. For example, a social planner would recognise that if the bank is less aggressive about exploiting its option to default on deposits it could raise deposits at a lower interest rate. So the bank and the saver might both reach different decisions were the feedback from risk-taking to lower interest rates was taken into account. But, the bank left to its own does not perceive anyway to commit not to gamble, so the feedback effect is ignored in its decisions.

Given these KTV assumptions, and various parametric assumptions about risk aversions, endowments, and the risk of investment project, the model can be solved for a set of allocations. KTV calibrate their parameters so that the baseline equilibrium has the saver making both deposits and equity investments in the bank. The bank lends to the entrepreneur and invests in some safe assets to cover early expected withdrawals (so that early loan liquidations can be avoided.) In the equilibrium, there is a positive probability of a run, because the bank does not fully recognise that its lending and borrowing choices affect the prices it pays. So it uses more deposit funding to exploit the default option that would be optimal if it recognised the effects on interest rates from doing so.

3 | Regulation

Regulation in the KTV model serves two purposes. First, as just described the agents do not fully internalise the effects of a run. So the private optimum will have an inefficiently high probability of a run. A run is destructive and lowers the welfare of all the agents: depositors do not get fully paid, the bank and the saver have their equity wiped out, and loans are liquidated early which harms the entrepreneur. Regulation to reduce the probability of a run can potentially be Pareto improving.

Second, as also explained, the limited liability assumption means that the borrower and the bank have an incentive to take excessive risk. The saver recognises these incentives in deciding how much to invest in the bank and the bank accounts for this when pricing its loan. But these price effects are not sufficient to overcome the distortion. Despite the higher interest rates, the limited liability assumption, on the margin, always leads to overinvestment. Provided that a bank-run has been controlled by other regulations, stopping the excessive risk-taking cannot be Pareto improving; the agents that are engaging in the gambling, will not want to be prevented from doing so.

The model thus embodies the two narratives that dominate the discussions of the crisis: excessive risk-taking and funding runs. Models that fail to respect our two principles are much more likely to exclude one of these possibilities. The purpose and effect of regulation are also likely to differ in models where eliminating distortions makes all agents better off, compared to models where eliminating distortions help some agents at the expense of others.

There are many tools that can be used to fix these two distortions. We briefly review five alternatives. The first one is raising banks' capital requirements. Higher capital requirements have multiple effects. The direct effect is to reduce the vulnerability of the bank to a run by reducing the reliance on deposit financing. Eliminating the run helps all three types of agents.

But, higher capital requirements indirectly exacerbate the limited liability problem. Higher capital requirements can only be achieved by when the bank induces the saver to buy more bank equity. When the saver buys more equity, the bank is offering...
less risk-sharing to saver. So this creates a powerful incentive for the banker to take more investment risk. Intuitively, substituting equity funding for deposit funding will reduce the required amount of liquid assets that the bank needs to hold, which frees up capacity to lend more. Quantitatively the direct effect of capital requirement tends to dominate the indirect effect so that the risk of a bank run still falls when capital requirements are increased.

A second tool is liquidity requirements. KTV model this regulation in the spirit of the liquidity coverage ratio which specifies the percentage of safe and risky assets. Left to themselves, banks only will hold safe assets to service their expected deposit outflows by the impatient savers whenever the cost of liquidating loans becomes very expensive. If the regulator forces banks to hold additional safe assets, the banks will endogenously respond by recognising that these assets allow them to raise more deposits. As the banks raise deposit funding and shrink equity financing, the benefits of the extra liquidity in preventing runs are reduced. The higher use of deposit finance also exacerbates the limited liability problems. Liquidity requirements of this type are less effective than capital regulation in overturning the distortions in the model.

A third possible regulation is deposit insurance. In the original Diamond and Dybvig model deposit insurance is a very attractive regulation because it wipes out the possibility of run without creating other problems. In the KTV model, deposit insurance does eliminate a run, but it severely encourages the banks to gamble. The banks load up on deposits at the expense of equity, and lend aggressively. Deposit insurance thus proves to be a two-edged tool.

A fourth potential regulation is the imposition of loan-to-value restrictions. This kind of regulation forces the entrepreneurs to consume less in period 1 and borrow less to finance their investment. This reduces the spillovers to the bank and the saver from a failed investment. The greater self-funding by the entrepreneur encourages the bank to fund its lending with deposits rather than equity. So there are competing effects for whether a LTV requirement will raise or lower the probability of a run: the reduced lending by the bank and higher LTV ratio makes a run less likely, but the desire by the bank to exploit the protection of limited liability leads it to use more deposit financing which makes a run more likely. The net effect on the probability of a run depends importantly on the parameter choices. Nonetheless, by crimping credit extension and having the banks offer less risk-sharing to the entrepreneurs, entrepreneurs are typically much worse off under this regulation in isolation.

Finally, it is possible to impose a tax on dividends paid by the bank on its profits. A dividend tax reduces the appeal of equity financing, so the bank and the saver want to shift towards funding loans with deposits. This leads to a reduction in interest rates on deposits which further helps the banks and hurts the savers. The increased use of deposit finance means the bank’s gambling is amplified, so that dividend taxes in isolation will not help reduce the risk of bank runs. This tool will benefit the bank at the expense of the savers.

From this brief tour of the various regulatory options we draw three important conclusions. First, accounting for the general equilibrium effects that arise from the endogenous choices of the agents in response to their incentives is important. The baseline equilibrium that obtains in the absence of any regulation has several subtle properties that emerge because of the forward looking behaviour of the agents. More importantly, only by recognising the countervailing choices that the agents will take once a regulation is imposed can we understand how they operate.

Second, once we have enough frictions in the model to allow the financial system to perform all three of its functions, no single regulation is enough to correct all the distortions. The problems associated with limited liability and runs are very different and we see that tools that fix one of them can often exacerbate the other.

Third, the ways that the different regulations operate, even to solve a similar distortion, can be very different, meaning that the consequences for different agents will vary. Generically, the bank’s gambling to exploit limited liability is usually bad for savers, so policies that reduce the bank’s use of deposit financing help the savers at the expense of the bank. Conversely, restrictions on the asset side of the bank’s balance sheet may not constrain the bank’s ability to gamble. So even though the risk of a run can be mitigated by working on either the asset or liability choices for a bank, the fallout from the policies will differ.
4 | CONCLUSIONS

The point of our brief tour of the KTV model is merely to show the rich and subtle ways that regulations can manifest themselves. The absence of any generally agreed upon principles, let alone the lack of a workhorse model, makes it easy to miss some of the interactions that we believe are important. We see the KTV model as a first step in a promising direction, rather than a complete framework that is ready to be used to quantitatively explore various regulatory options.

Nonetheless, the KTV model already leads us to three robust conclusions. First, in models that are rich enough to realistically capture the different roles of the financial system, the different roles are likely to be justified by a variety of frictions. Put differently, it is well understood that the form of financing for a firm only matters when the assumptions underlying the Modigliani and Miller capital structure irrelevance propositions fail. We think that to explain the various contributions of the financial system to supporting economic activity, several of these assumptions must fail. Our cursory exploration of the KTV model suggests that correcting all of the distortions caused by the frictions will require multiple regulatory tools.

Second, using the limited set of tools which we described to attack the various distortions have very different allocational effects. One of the assumptions in the KTV model is that markets are not “complete”, meaning that complete hedging is impossible and default is a possible outcome. In that case, there are not going to be market based schemes (and associated prices) to evaluate the full social costs of transfers of resources from one agent to another. A social planner in these circumstances will have to decide how much weight to place on various agents to evaluate policies. This makes it much harder to rank and compare alternative policies.2

Finally, the fact that competing policies have such potentially different allocational effects creates strong incentives for the disadvantaged parties to try to evade the regulations. One prominent feature of the last crisis was the role that regulatory arbitrage played in making the risks in the financial system more opaque and vulnerable. The KTV framework does not allow for regulatory arbitrage, nor do most other models that are used in discussions of macroprudential regulation. This is an important defect of these models and a critical area for future research.

---

2 This is a much more general issue. Welfare analysis in models with incomplete markets is not straightforward.
REFERENCES

Admati (A. R.), DeMarzo (P. M.), Hellwig (M. F.) and Pfleiderer (P.) (2010)
"Fallacies, irrelevant facts, and myths in the discussion of capital regulation: why bank equity is not expensive", Rock Center for Corporate Governance at Stanford University, Working Paper, No. 86.

Allen (F.) and Gale (D.) (1997)

Benston (G. J.) and Smith (C. W.) (1976)

Calomiris (C. W.) and Kahn (C. M.) (1991)

Cass (D.) and Shell (K.) (1983)


Diamond (D. W.) and Dybvig (P. H.) (1983)

Diamond (D. W.) and Rajan (R. G.) (2001)

Goldstein (I.) and Pauzner (A.) (2005)

Kashyap (A. K.), Rajan (R. G.) and Stein (J. C.) (2002)

Kashyap (A. K.), Tsomocos (D. P.) and Vardoulakis (A.) (2014)
"How does macroprudential regulation change bank credit supply?", University of Chicago, mimeo.

Kotlikoff (L. J.) (2010)
Jimmy Stewart is dead, New York: John Wiley & Sons.

Morris (S.) and Shin (H. S.) (1998)

Rosengren (E. S.) (2011)