Fundamental principles of financial regulation and supervision

Jan A. Kregel, Mario Tonveronachi
Fundamental principles of financial regulation and supervision

Jan A. Kregel, Mario Tonveronachi
Tallinn University of Technology, Estonia

Abstract: The financial system is a private-public partnership coming from government ceding the right to produce means of payment with the related permission on leveraged lending services, against the acceptance of rules designed to ensure stability for both individual institutions and the financial system. The experience shows that market-based regulation does not produce the wanted results, while rules-based and principle-based regulatory systems are prone to regulatory avoidance and capture, especially with complex regulatory schemes. While the reaction to the recent crisis has prompted a wide range of financial reforms, in a duel to match complexity with complexity, the previous approach based on leaving market forces to mould the financial structure with few if any constraints maintained. The paper shows that this approach adopts faulty or casuistic policy implications derived from both the laissez faire and the second-best versions of mainstream economic theory. However, some of its basic features, such as regulating institutions and products and not functions, and as promoting the international level playing field, are not coherent with its reputed theoretical foundations. Furthermore, the absence of strong principles and the impossibility to derive conclusive quantitative proposals from cost-benefit evaluations leaves an unacceptably wide area of discretion for experimentation with trial and error processes, easily leading to weak or distorted regulation. The difficulties experienced within this framework to deal with problems such as those posed by systemic institutions, shadow banking, weak rules and supervision, distorted risk evaluation, high compliance costs, etc. has convinced some observers that a ‘revolution’ in economic thinking and policy is required. Following Minsky, the conclusions review a heterodox approach to financial fragility and regulation. Characterising banking in terms of
liquidity creation through acceptance, and distinguishing it from the production of liquidity by other financial institutions, it concludes that financial organisations should be regulated according to their function in providing liquidity of different types to the financial system.

**Key words:** Financial regulation, financial supervision, financial fragility, Hyman Minsky.

**Date of publication as FESSUD Working Paper:** March, 2014

**Journal of Economic Literature classification:** G01, G21, G24, G28, G32

**Contact details:** Jan A. Kregel, jankregel@yahoo.com; Mario Tonveronachi, tonveronachi@gmail.com. Tallinn University of Technology, Ragnar Nurkse School of Innovation and Governance, Akadeemia tee 3, 12618 Tallinn.

**Acknowledgments:**
The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement nº 266800.

**Website:** [www.fessud.eu](http://www.fessud.eu)
1. Introduction

Prudential financial regulation emerges from the fact that financial systems have always been a private-public partnership. Governments generally retain the monopoly right to produce coin and currency. They often cede this right to private sector financial institutions against a monetary payment or a set of rules and regulation of conduct. Thus, the private sector acquires the ability to offer and compete against government means of payment by accepting rules and regulations that govern their behaviour. This is the basis of the private-public partnership that has characterized every monetary and banking system since the creation of the Bank of England. The government regulation that has accompanied permission to the private sector to offer means of payment and leveraged lending services to the private sector has been designed to ensure the stability of the private financial system as well as to ensure that it provide support for the general policy objective of growth in output and employment. The regulation that seeks to limit the operations of private financial institutions in order to ensure the stability of individual financial institutions is generally denominated “prudential” regulation. The regulations that seek to ensure that the operations of the private financial system do not lead to system instability are generally denominated “macroprudential” regulation. Traditionally bank regulation has concentrated on the stability of individual banking institutions, but increasingly attention has been focused on the systemic impact of financial institutions behaviour and the need to provide regulation that looks beyond the individual financial institutions. The paper outline the initial approach to the stability of individual financial institutions, its extension within the European context as well as the logic of systemic instability and regulations to meet it.

2. The Foundation of Government Regulation of Banking

All Sovereigns, whether Princes, Dictatorships, or Constitutional States or Monarchies, have usually reserved to themselves the prerogative to regulate trade, production and the issue of means of payment in the form of currency and coin. Free voluntary exchange in public markets, which is now considered the basis of economic activity, was initially regulated by the decision of the Sovereign. These regulations applied to the particular
locations in which markets could be held, as well as to permissions for particular dates or special occasions. It was also common for the Sovereign to cede such rights to private individuals, as in the “farming” (granting of the right to collection) of taxes and to duties on trade. It was also common to grant rights to provide alternatives to official coin and currency.

The best known example of the latter is the granting of a Royal Charter providing for the creation of the Bank of England and the subsequent interpretation of the Charter as providing the Bank with the monopoly right to the issue of bank notes. It was in the specification of the conditions of the grant or charter under which these activities could be carried out that the regulation of financial institutions by the State was born. Thus, there can be no question of the legal source of the power of the State to regulate financial institutions; there remains the question of the extent to which the State should exercise this power and on the ability to make regulation effective.

3. The Impact of Regulation

The formal permissions given to the private sector to engage in activities that are the province of the State represent a limitation on the extent of this activity. These regulations generally create an incentive to evade or avoid them. This can take several forms. The first is simple evasion or the flaunting of regulation due to inability of the State to effectively enforce regulations.

The second is avoidance of regulation through the implementation of activity that is not technically in violation of existing regulations. An example of the latter is the operation of US banks in the London Eurodollar markets where they were able to engage in activities that they were forbidden in their domestic markets by their national regulators.

The third is financial innovation; the creation of activities similar in scope and intent, but which were not expressly covered by existing regulations. Again, the Bank of England provides the best example. While its Charter was interpreted as granting it a monopoly on the issue of bank notes that served as a substitute means of payment to officially minted
coin, the Charter did not cover, and therefore did not regulate, the conditions on which other financial institutions could create deposit accounts subject to cheque and provided an effective substitute to Bank of England notes as a means of payment. The pervasive nature of incentives to circumvent regulations and the existence of competitive innovation in the financial system means that regulation must be continually revised and updated to keep pace with the evolution of the organisation and operation of the financial system if it is to be effective.

4. The Form of Regulation

The form that regulation has taken has varied widely over time and over different structures of the financial system. Regulations which exclude specific types of financial operations are subject to erosion by innovations that do not fit the exclusionary list precisely. The same is true of inclusive regulation which limits to a specific list of accepted activities. Similar problems arise in both rules-based regulatory systems and principles-based regulatory systems. The former, which promulgates specific rules to be applied by supervisors, may be subverted by innovations that fall outside the set of rules. The second, which is based on a set of principles to be applied, and thus designed to allow human cognition to identify and avoid the problems of the blind application of rules, leaves a degree of discretion in application by the supervisor which often leads to excessively lenient application of principles or to what has come to be called “regulatory capture” in which the supervisor absorbs the objectives of the institutions to be regulated.

Finally, throughout the history, the States have experimented with the use of the actions of competition in free markets as the basic form of regulation. An example is the era of free banking in the US when restrictions on the granting of bank charters were eliminated. It was argued that restrictions on entry into banking were not needed because the market mechanism would weed out inefficient or fraudulent bank operations. The result was that, in the absence of reliable information concerning the condition of banks, rumours real or imaginary concerning the ability of a bank to meet its outstanding note issue led to
contagious bank runs that eventually spread to the entire system and produced financial crisis.

An alternate form of market-based regulation is counter-party regulation or self-regulation. Here the argument is that it is in the interest of individual banks to evaluate the condition of the other banks with which it deals, thus providing an automatic policing of fraudulent and/or excessively risky activities. The argument is that in their own self-interest bank management and shareholders would not permit risky or fraudulent activity that would threaten their own existence and livelihood.

The experience of deregulation and liberalisation of financial markets, removing the monopoly on deposit-taking to insured commercial banks and allowing market competition in the provision of banking services and the setting of bank interest rates in the US from 1980 to 2008 has brought both of these forms of market-based regulation into question. Indeed, one of its foremost proponents, Alan Greenspan has admitted that private self-interest in private financial markets was not capable of providing effective self or counter-party regulation of risky or even fraudulent activities.

5. The Costs of Regulation

Recently there has been increased emphasis on the efficiency of financial regulation. This is because the banking sector, in particular after the recent financial crisis, is currently considered as being subject to excessive regulation. Measuring regulation in terms of pages or regulatory submissions of information, this is undoubtedly true. This suggests that the quantity, cost and effectiveness of rules do not necessarily coincide. The New Deal financial regulations that were in place for over fifty years covered less than 40 pages, while the recent Dodd-Frank Act comprises over 800 without consideration of the specific additional reports, rules and institutions that is requires to be created. History of financial systems shows that financial crises have accompanied the capitalist system since its early development. However, the increasing frequency and seriousness of financial crises in the last decades despite more extensive and pervasive regulation suggests that a rethinking of
this approach is overdue and that more simplified regulations may be more effective than increasingly casuistic writing of regulation.

6. The Motivations for Regulatory Changes

History also shows that major regulatory changes have been motivated by the real costs in terms of loss of output, employment and accumulated wealth of financial crises. Although it is often possible to identify the historical rationale for single regulatory measures in specific crises, the accumulation of such ad hoc regulatory responses has meant that it is not easy to find a consistent and clear set of principles behind actual regulatory and supervisory frameworks.

Several factors are responsible for this. Financial legislation is normally the result of compromises between policy alternatives and diverging specific interest groups or ideologies.¹ Specialised national or international regulatory bodies may not necessarily share a common vision, or may find it difficult to reconcile divergent visions of the scope and objectives for the financial system. As noted above, a changing regulatory environment driven by both internal and external financial conditions creates an ever-present threat to regulatory effectiveness, and modifications to meet specific circumstances may be introduced without regard for their general consistency on either a domestic or an international level. In addition, if regulation is primarily based on principles, the distinction between regulators and supervisors is compromised as the latter use their discretionary powers to adapt their practices to what they perceive as being appropriate to new circumstances.

7. Changes in Regulation in Response to the Current Crisis

The recent crisis has spurred a wide revision of financial regulation, whose breadth might be perceived as a serious attempt to give it consistency and effectiveness. However, this re-regulation process has not been the result of a radical rethinking of the existing set of

¹ See for example the debate promoted by the so-called Economic Theory of Regulation that concerns how much regulation is pushed by social interests in solving market failures or by the pressure of interest groups on politicians. Cfr. Stigler 1971, Posner 1974 and Peltzman et al 1989.
regulations or the overall approach to financial regulation. According to the G20, an international group that has taken the lead in proposing a new international regulatory structure, the crisis was the result of ‘excesses’ that were produced by specific lacunae in the previous regulatory framework. On this view, all that has to be done is to deal with specific areas.

As the process of regulatory revision has proceeded, it has been complicated by an unexpected polarisation of position, even among those regulators, academics, analysts and industry representatives that are thought not to favour a radical departure from the previous approach. As an example, resistance to implementation of the new Basel III Accord has become widespread, although this has been justified by a number of very different objections. As first principles are increasingly called on to adjudicate divergent positions and settle the disputes, this background document attempts to clarify the theoretical foundations of alternative regulatory approaches.

8. Mainstream foundations of financial regulation

The fact that the multiple decisions of a myriad of independent individuals dealing in free markets could produce sustained increases in wealth and well-being led economists to seek the mechanism of coordination for a self-adjusting system. The answer was the operation of the price mechanism in free competitive markets. This is the essence of mainstream theory and the foundation of the general equilibrium model, which provides its formal reference.

The problem with this model, even in its more advanced Arrow–Debreu version, is that it cannot accommodate money and finance in any meaningful way without losing its distinctive outcome of a unique Pareto optimum equilibrium.2 As demonstrated in the capital controversy spurred by the work of Piero Sraffa, well-behaved parables cannot conceal inherent specification errors.3 Nevertheless, mainstream theory treats real

---

2 This even maintaining the plethora of other restrictive assumptions.

3 For a critical analysis of mainstream economics see Roncaglia 2005.
problems and policy questions by proceeding “as if” a consistent general equilibrium model were capable of validating simple parables based on partial application of the theory.

The fact that experience never conforms to the results of an optimum self-regulating system is explained in terms of ‘imperfections’ that impede the real world system from attaining the ideal Pareto conditions. From the basis of this ideal theoretical system, there are two different approaches in the way in which these imperfections are incorporated in the analysis and in the implications drawn in terms of policy.

In the more orthodox approach, the role of public authorities should be limited to ensure the full implementation and application of property rights and to eliminate any potential obstacles to the free working of the price mechanism. This is the essence of a laissez faire capitalist system. Specific imperfections should be removed and, where this is not possible, dealt with in a way that moves performance closer to the ideal system.

The alternative perspective is grounded in the theory of second best which allows the introduction of additional imperfections when initially existing imperfections cannot be eliminated. As a clarification example, operational efficiency is one of the conditions required by the laissez faire model. However, removing operational inefficiency may appear to be welfare enhancing unless there are allocational inefficiency present that defy elimination. In this case, a second best solution could require introducing additional inefficiency in the form of increasing transaction costs.4

Despite the common reference to the general equilibrium model, both approaches are partial and tend to look at single departures from the optimal equilibrium state, so that the discussion of regulation produces intervention in the operation of the system that is micro and piecemeal, rather than systemic.

Imperfections, and potential public intervention to remedy them, concern three interrelated dimensions: the scaffolding of the private financial system, specific price mechanisms and the public regulatory and supervisory institutional framework.

According to the laissez faire approach, the structure of the financial system endogenously adapts to the needs of the economy. Financial innovations, in both private institutions and financial products, are the main drivers of this adaptive efficiency.\(^5\) The “efficient market hypothesis” presumes agents’ best use of available information in allocating funds, independently of perfect information and complete markets. Arbitrage is the expression of competitive forces and the necessary condition for price discovery leading to equilibrium. This explains the focus on eliminating imperfections and barriers to the free movement of capital, i.e. on operational efficiency. Although the laissez faire equilibrium may not be optimal, public intervention is viewed with suspicion because it may introduce more distortions than it is designed to remove.\(^6\) In this vein, the approach favours studies on the distortions of political processes and decisions concerning economic matters, decidedly rejecting the possibility of politicians acting as benign planners. Specific legal implications follow, generally tending to restrict the scope of public intervention and to tie the hands of politicians and bureaucrats with fixed rules. In any case, when specific public action is clearly directed at removing, or at least lessening single departures from optimal conditions, they are accepted without question. When public action takes a different direction, the onus of proof is shifted from the proponents of laissez faire to public authorities who must assess the net benefits to be produced from their action.\(^7\) This is a crucial passage, since to be effective it requires grounding the proof on objective or clearly verifiable quantification of the net benefit. It should be noted that the implementation of

---

\(^5\) “Financial functions are more stable than financial institutions – that is, functions change less over time and vary less across borders. Institutional form follows function – that is, innovation and competition among institutions ultimately result in greater efficiency in the performance of financial system functions”. Merton and Bodie 1995.

\(^6\) Economists’ own inability at successfully modeling the system along the previous lines may play a part in their disbelief in public intervention. Hence, markets know better, i.e. the capitulation to an unproven invisible hand.

\(^7\) Allen and Gale [2003, p.6] offer a clear example of this approach: “Under standard conditions, the incidence of financial crises may be socially optimal in a laissez-faire system. [...] And if not, for example, if financial crises involve deadweight losses, it should be recognized that regulation also involves administrative costs
many of the mandated rules under Dodd-Frank have been suspended or delayed because the regulatory agencies have not been fully compliant in providing assessments of the costs and benefits of the proposed rules.

In order to clarify this point let us refer to the concept of negative externality, which is often used with reference to possible costs coming from unregulated markets. Pigou [1920] formalised the concept of externalities in terms of prices that do not fully reflect the social costs (or benefits) associated with a given activity. A tax (or a subsidy) could align social and private costs by adjusting demand and production through higher (lower) prices. Taxing air pollution is the standard example.\(^8\) However, moving from a general theoretical proposition to a specific policy measure requires being able to perform a series of calculations. Roughly, it is necessary to:

1 - Measure the level of the externality
2 - Identify the causes of the externality in terms of variables that can unambiguously be transformed into quantitative indicators
3 - Quantify the relation between the externality and its causes
4 - Quantify the relation between the regulatory tools and those causes
5 - Quantify the costs and unintended ‘distortions’ coming from regulation
6 - Quantify the social cost-benefit frontier

Especially point 6 makes these calculations a political and not a technical exercise, which conflicts with the regulatory goal of giving stable rules to private markets. In addition, as suggested by the classical technologically based air pollution externality, these calculations encounter greater difficulty in the financial sphere due to its specific features and the more evident limits of partial analysis. For example, it could be difficult to directly associate

\(^8\) Here reference to the Pigouvian tax is made just to illustrate the problematical aspects of a cost-benefit analysis applied to financial externalities. The question of whether the management of systemic risk is preferable via some form of taxation or the existing regulation based on capitalisation is not considered here (Coulter, Mayer and Vickers, 2013).
externalities to a given product. A few million junk bonds dispersed in world portfolios produce positive externalities, while hundreds of billions concentrated in few hands constitute a serious threat to financial stability. Furthermore, the shape of the relation between junk bond pricing and their potential damages depends on a myriad of other financial and economic features, thus making it impossible to precisely define their relation, which would in any case be subject to large unforeseeable fluctuations.

The *laissez faire* framework of reference and the stalemate that the net benefit rule produces on public policies render this approach *de facto* a theory of de-regulation.

The second best approach gives wider and stronger relevance to imperfections and public intervention; it is thus more policy friendly. An array of imperfections such as asymmetric and costly information, incomplete contracts and markets, transaction costs, non-competitive markets, conflicts of interest, etc. are considered as impediments to constrained Pareto efficient results, so that public intervention may enhance social welfare. This introduction of imperfections as physiological features in the orthodox model gives a Keynesian interventionist flavour to the results. However, imperfections do not change the fundamental nature of the approach that wherever possible, policy should be oriented toward their removal. Otherwise, public policy should be directed towards creating incentives guiding private agents towards second best solutions.

For the specification of the structure of the financial system the starting point is similar, with endogenous forces leading to adaptations to the needs of the economy. The difference is that fully integrated imperfections produce agency problems and distorted incentives in risk management, affecting both financial products and the nature and governance of

---

9 The Coase solution of eliminating externalities through private bargaining is impeded in the financial sphere by the plurality of parties that make clear definition of property rights difficult. For a compact analysis of imperfections in the perspective of public intervention, see Stiglitz 1994.

10 As Rodrik [2013] puts it, conventional wisdom is questioned from within rather than from outside.

11 The danger of a piecemeal approach to regulation and supervision in front of a plurality of imperfections is to produce unwanted mutual interactions between ad hoc measures. See Goodhart et al 2013.
private institutions. The nature of negative externalities makes public regulation and supervision imperative. For example, if asymmetric information and lack of proper incentives prevent creditors from monitoring banks, regulation steps in to mimic risk-averse creditors and avoid excessive risk taking. Consistent with the general theoretical model, regulation should modify, wherever possible, market incentives to guide operators towards ‘correct’ risk pricing. However, if incentives do not work as expected, further ‘imperfections’ may be introduced, for example with structural measures disconnecting areas of bank business.

Incomplete contracts and agency problems also characterise regulation and supervision. Fixed rules and monitoring and arbitration activities cannot be perfect, allowing for significant degrees of discretion. Special attention must then be given to the design of both institutions and rules. Political distortions, which now accompany equally serious private ones, do not impede public intervention, but require that a carefully designed institutional framework produce an acceptable equilibrium between rules and discretion. This has led in recent decades to a preference for shifting responsibility to independent technical authorities, making them subject to specific legislative mandates. However, being imperfect rules, mandates do not eliminate discretion and political ‘interference’ by authorities lacking democratic representation, thus leaving open the question of their control and the establishment of a set of incentives consistent with their mandate.

A relevant question for both approaches is whether regulation should concern institutions, products or functions. Adaptive approaches that leave financial structure endogenously determined by financial innovation are not consistent with direct regulation of specific institutions and products. If it did, it would support the creation of innovations for regulatory avoidance. An alternative is to follow the functional approach favoured, inter alia, by Merton and Bodie (1995). The authors make the case of banks using insured deposits to fund loans. They argue that the function carried out by banks in the payment system benefits from insured transactions deposits because monitoring costs are thus held at a minimum. On the contrary, using insured deposits to fund loans produces moral hazard. Regulation
should then distinguish between the two functions, obliging banks to fully collateralise insured deposits with Treasury bills and fund loans with capital and other forms of debt. This would not require a narrow bank structure and would permit other financial intermediaries, such as mutual funds, to accept insured deposits under the same regulatory conditions. A similar approach has been proposed by James Tobin [1987] and Hyman Minsky [1995].

Functional regulation should also permeate the institutional structure of supervision. After singling out the functions to regulate, those that can be usefully grouped together could be consolidated under a single authority, which should leave different private institutions free to perform similar functions. This is not the same thing as specialising supervision according to general objectives, as done by the two-pillar approach that envisages a prudential supervisory authority and a consumers’ protection authority.

The functionalist approach does not require specification of the nature of regulation. Once the different functions are singled out, it is, for example, equally consistent with self-regulation or invasive public action as in the previous example. In other words, regulation for any function may incorporate market standards or policy goals. However, a fresh look with a functionalist lens might have favoured the recognition of the push given to the dominance of the financial sector by the introduction of contractual savings and pension plans and the associated extensive professional wealth management, and to its potential conflicts with the lending function on which the literature tends to focus.

In any case, adopting the functionalist approach inside the mainstream framework does not dispense with the obligation to compute the net benefit of each regulatory measure. Although the second best approach may make the case for intervention in general terms

---

12 The next section will return to the functional approach to regulation because its methodology can be made consistent with non-mainstream theories.

stronger, it still requires cost-benefit evaluations, whose results are far from certain having to apply margins of discretion to evaluation of trade-offs.

Perceiving the impossibility of giving conclusive quantitative answers, some regulators have proposed a sort of trial and error process, finally leading to an acceptable calibration.\textsuperscript{14} This is almost an admission of working on weak fundamentals and it is a dangerous request since it would not be excessively unfair to state that the trial and errors of the last decades of deregulation and prudential re-regulation have produced a string of increasingly severe systemic crises. The recent debate in Europe on the Tobin tax is a further example. Its quantitative formulation is entirely tentative, much oriented to produce the financial resources necessary to subsidise some sort of financial or social safety net, with overall net effects open to widely different evaluations. A trial and error process is admissible only when reliance on strong principles reduces its scope.

As repeatedly stated, both approaches to public intervention are directed at improving the price mechanism. As already seen, a wider range of public intervention is, however, consistent with the second best approach in which the price mechanism may not work in normal circumstances, such as credit rationing. In the presence of some strong signal, asymmetric information is capable of making markets disappear. It may not be possible for public authorities to promptly monitor strategic regulatory variables, or monitoring might be excessively expensive. The solution of conflicts of interest may be impervious to prudential rules (e.g. control of banks by non-financial firms and vice versa). Hence, structural measures, such as forcing the bank-industry separation or the specialisation between commercial and investment banking, may be consistent with this approach, although as measures of last resort.

Structural measures coming from the second best approach remain, however, part of the choice between rules and discretion applied casuistically, and do not come from an

\textsuperscript{14} The various releases of the Basel Accord represent the most visible example of a trial and error process.
alternative view of basic economic principles, of the role of the state and of financial morphology.

9. Does current regulation meet mainstream fundamental principles?

As noted above there are several reasons why actual regulatory regimes do not necessarily follow from a set of consistent fundamental first principles. It should be added that such departures might also measure the degree of discomfort with some of those principles. As illustration, consider the phenomena of shadow banking, the two pillars of the current regulatory framework - the international level playing field and the Basel Accord on banking regulation and supervision, and the debate on institutions that are too systematic to fail.

The explosion of the shadow banking system well illustrates the contradiction between the laissez faire approach to the evolution of the financial structure and the partial approach followed by regulation. The latter necessarily leads to regulation of institutions, leaving innovation to mould institutions and products as a response to regulatory costs and constraints. The problem is not just that regulation is obliged to adapt; running after innovations with a delay, regulation and supervision become sources of distorted incentives for significant periods. For instance, while in the pre-crisis period regulation was concentrating on bank capital, the critical build-up of fragility was outside banks via the interaction between funding and market risks. How much the laissez faire approach is rooted in mainstream regulation is shown by reiterated affirmations that the current efforts to re-regulation must not create obstacles to private sector financial innovations. Even the current efforts directed at regulating institutions that are part of the shadow system may easily result in new techniques of regulatory avoidance. This provides substantiation for the adoption of a functionalist approach to regulation and supervision.

The idea of creating a regulatory level playing field accompanies the liberalization financial flows and institutions in creating a global space for competitive arbitrage. The underlying assumption is that the same set of rules and principles can efficiently govern different local situations, whose idiosyncrasies can be accommodated within the general framework. To
this end, regulatory standards should consistently come from basic principles common to every market economy. Following the orthodox approach, one should consider the question of why there should be any common regulation at all, save when necessary to dismantle barriers. If a good theoretical case can be made for a specific regulation, the test of the net benefit must be passed and there is no reason why we should obtain the same result in every country. In other words, the orthodox approach surely leads to a de-regulated level playing field regarding obstacles to arbitrage; as to specific measures, it cannot justify the level playing field as the result of an active regulatory process.

The problems facing the second best approach are different since it may question both complete liberalization and the perfectly level playing field. The terms “complete” and “perfect” stand to indicate partial exceptions to the same basic general scheme. Some “sand” may be necessary to slow the gears of the arbitrage machine, as some common rules may be required at different level in different countries. The fact is that the same theoretical foundations of the second best approach mitigate against postulating and effectively reaching the level playing field. The relevant imperfections that justify regulation may vary widely across countries, posing different priorities. Discretion coming from a few imperfect rules and many principles can be used to push finance in different local directions under an apparently common regulatory framework. Besides, absent an international supervisor with enforcement powers, these standards may only take the form of soft laws, which further shifts the balance from rules to discretion. The question then arises in both approaches: how to justify the merit of international standards with respect to national ones. If the answer is to let global financial institutions save on regulatory costs, it is difficult to justify on the basis of first principles.

The trial and error path followed by the various releases of the Basel Accord shows the enormous difficulties in embedding the level playing field into a common regulatory and supervisory framework based on mainstream principles. The problems that prompted the first 1975 Basel Concordat on the co-operation between national authorities in the supervision of banks’ foreign establishments stemmed from the explosion of financial
globalisation and have remained substantially unresolved. The problems related to how to supervise global banks and how to deal with their eventual failure remain. In principle, the Basel Accord would have helped the Concordat by laying down some minimum common standards for internationally active banks. Basel I was explicitly directed only at global banks. As was to be expected, international co-operation for bank resolution encountered formidable obstacles and has remained on the wish list. Renewed post-crisis attempts continue to show unaltered difficulties. Despite general discourses favouring co-operation, the Fed proposal to oblige foreign banks operating in the USA to hold capital and liquidity locally is implicitly directed at creating a national, not international, level playing field for bank resolution.

In a global environment with unrestricted banking activity, the attempt to mitigate the weight of resolution procedures should have required global banks to carry additional buffers to withstand unexpected shocks. Merton (1995, p. 464) explains why capital was chosen as the buffer.

The management of risk has traditionally focused on capital. Equity capital is the ‘cushion’ for absorbing risks of the institution. It is a wonderful, all-purpose cushion. Why? Because management need not know what the source of the unanticipated loss is. They do not have to predict the source of the loss, because equity protects the firm against all form of risk; it is in that sense an all-purpose cushion and thus it is very attractive for managing risk. As we all know, equity capital also can be quite expensive for exactly that reason. One can formally employ theories of agency costs, taxation and so forth to supply reasons why equity financing can be expensive.

Furthermore, linking expensive capital to risks should also produce incentives towards avoiding excessive risk-taking.

The fact that regulatory authorities judge that unregulated banks normally fail to build up enough capital means, as the above quote clarifies, that they believe that relevant and
unavoidable imperfections impede the working of the Modigliani-Miller theorem. In addition, banks underestimate the negative externalities coming from unexpected shocks. Regulators thus solve the Pigouvian problem by linking the capital ‘tax’ to the potential cost of the externality, which is ultimately borne by the taxpayer. Since for global banks the difficulty is the sharing of losses among taxpayers of different jurisdictions, the capital cushion should be set at high levels. In other words, it is the speed and efficiency of resolution procedures that should determine the amount of capitalisation. Note that bank failures are easily managed with respect to nationally-based small and medium sized banks. The problem comes from large and complex global banks facing several national regulatory authorities. This may suggest the presence of an implicit contract. The bank profits from being global but has to pay a tax high enough to shield different national taxpayers. What is hard to understand is why the capital tax should be extended to smaller banks, whose different structure of activity should call for a different type of regulation.\textsuperscript{15}

The ability of the buffer to shield taxpayers depend on three critical questions: the definition of capital, the level of capital requirement, and the correct measure of the risks. The history of the various releases of the Basel Accord turns around these questions. The history of financial crises in the last three decades strongly suggests that something went astray.

The definition of regulatory capital according to its ability to absorb losses is not independent of the legal code of each jurisdiction, especially when hybrid instruments are allowed. For a regulation based on capitalisation it might seem odd to find, after two decades from its first implementation, wide international differences in the instruments composing the various layers of regulatory capital. Although conscious of these difficulties, the BCBS has again laid down a series of principles regarding the instruments eligible for inclusion in Tier 1 capital for the implementation of Basel III. The dissatisfaction in respect of the BCBS (2012) with the new EU CR Directive and Regulation is just an example of the

\textsuperscript{15} Going from Basel I to the later releases we may note that the progressively looser wording in the exclusive reference to internationally active banks. They are now spelled out as subjects to stricter requirements.
obstacles that the implementation of a general model encounters when facing the structural differences in different financial systems.

The minimum total capital requirement, initially at 8% of risk-weighted assets, appears to be just the capital international banks held on average at the time of Basel I. This means that the real goal was to stop the long-term decline of capitalisation, not to reverse it. It is then not strange that later experience has shown that that level was not a safe buffer against bank crises, even with the later addition of market and operational risks to the denominator. Especially after the implementation of Basel I.5 and II, sensible measures such as straight leverage show a continuing decline of capitalisation. It is doubtful that the much-advertised recapitalisation coming from Basel III will make banks return to positions similar to what they had before the first release of Basel. Indeed, it is only with Basel III that a super-buffer related to the relevance as global systemically important banks G-SIB appears, which our previous narrative justifies in terms of increasing resolution difficulties.

Disregarding the fact that the level of capital also depends on the denominator, after the publication of the new more stringent Basel III requirements, a lively discussion started between those expecting that it would generate large negative effects on GDP growth and those that consider the new requirements insufficient. In large part, the discussion may be reduced to whether banking approximates the Modigliani-Miller theorem.\(^{16}\) The less the cost of funding depends on its composition, the lower would be the effects of higher capitalisation on bank profitability and finally on the cost of credit, thus lowering social cost with respect to the benefits coming from expected higher stability. We are back to the cost-benefit analysis, where the divergent opinions come from giving different weights to the two terms of the equation. The fact is that neither side is able to produce indisputable figures to support a general statement. This does not come as a surprise. As far as the analysis concerns national banks, it should consider the structural specificities of each country. The figures must be country-specific since the evaluation concerns the effect of

\(^{16}\) Not even the more ardent supporters of very low leverages assume that the M-M theorem applies perfectly to banks.
capital requirements and crises on GDP, the degree of financial competition inside and outside the banking sector, the degree of “bankarisation”, the efficiency of capital markets, and so on. As for global banks, a general statement may only result from a common reference, but one must not confuse globalisation with homogenisation. It seems, again, that a complex reality is dealt with weak principles that permit excessively diverse and hardly reconcilable regulatory proposals.\(^\text{17}\)

It might be thought that even if the absolute level of capitalisation cannot be defined, it would be possible to deal with the diverse national conditions by means of the level playing field and common levels of capitalisation by shifting all the weight to the denominator, i.e. on risk weighting. This is why the supporters of Basel resist the adoption of straight gross leverage ratios. This raises more serious problems that have produced increasing dissatisfaction among supervisors on the reliability of the adopted methodologies.

Recall the implications of basing capital regulations on regulatory risk pricing. As already noted, if the Modigliani-Miller theorem does not hold, banks’ decisions on leverage depend on how they (and sometimes their creditors) evaluate risks. Regulators take lessons from experience in judging that banks and markets fail to price the costs that will eventually result from a bank failure. Before Basel III, the evaluation of these costs was limited to the micro-level, leaving general policy preconditions to shield banks from macro shocks. Basel III adds to an unchanged risk methodology a countercyclical capital buffer to smooth changes in banks’ activities over the economic cycle.\(^\text{18}\) In principle, an averaging over the cycle of this buffer should be near zero. Dynamically, this means that changes in risk mispricing are due to overreactions to exogenous factors and are not linked to the endogenous dynamics of financial structure. As we shall see in the next section, this is a crucial aspect on which to build a non-mainstream approach to finance and regulation.

---

\(^\text{17}\) Another argument could be that since we are not certain of the denominator, the risk weighting, it is better to adopt a high ratio. However, the problem does not change even if we cancel risk weighting in favour of strict leverage ratios; it is still necessary to decide on their level.

\(^\text{18}\) A 2.5% conservation buffer is also added to the 8% requirement to promote prompt corrective actions by banks hit by idiosyncratic shocks.
Apart from these reactions, the goal as expressed by the BCBS is to align regulatory capital with economic capital, which means aligning regulatory risk pricing methodologies to the best available industry practices. Since Basel II, this has meant to expand the adoption of Value at Risk (VaR) methodology to the risks considered under Pillar 1. Disregarding the many theoretical and empirical deficiencies of this method and the wide differences observed on how banks price the same portfolio, it is quite sensible to believe that the private sector uses the best quantitative methodology that it can develop and afford, although complementing it with qualitative assessments and with an awareness of its deficiencies. With the right incentives, private operators are in the best position to decide which methodologies to adopt, how to mix quantitative and qualitative evaluations, when it is suitable to introduce modifications in their methodologies, etc. The problem is why regulators should themselves adopt systemically unreliable methodologies and ossify them as regulatory standards, knowing that in so doing they amplify distortions and distort incentives. The new fashion of applying stress tests to compute an addition to minimum capitalisation necessary to withstand shocks does not change the matter since the reference point, the minimum capital, comes from those methodologies.¹⁹

This risk framework is hardly consistent with the two approaches here presented. According to the orthodox approach, the solution should be to give the appropriate incentives to private decisions, and Basel is not capable of doing so. According to the second best approach, the failures that produce private mispricing should equally apply to regulators and supervisors. Moreover, the enormous discretionary powers given to supervisors hardly reconciles with the problems afflicting their governance.

A subtle logical inconsistency is also produced by the coupling of Basel with the principles of bank resolution sponsored by the FSB that are being adopted in several jurisdictions. According to Basel, supervisors must certify the risk methodologies adopted by each bank,

---

¹⁹ The design of stress tests may permit the inclusion of specific features and to adopt scenarios proportional to the systemic relevance of the intermediary. In practice, however, it appears quite impossible to apply them to global banks with thousands of foreign dependencies. Again, a regulation designed for large banks may be effective only for smaller ones, for which it is not said that they are the most appropriate.
stress tests included. Once the supervisory review process of a bank ends with full satisfaction of supervisors, *de facto* they are giving a guarantee on the resilience of the institution. If a bank were to fail in circumstances that are within the validated behaviours and parameters, that guarantee should consistently result in, at a minimum, sharing the costs of the resolution between the private and the public. On the contrary, while increasing the regulatory tax and the powers of supervisors for both discretion and enforcement, the current trend is to promote resolution procedures that make shareholders and creditors fully liable to bank losses, while shielding supervisors (and the state) from *bona fide* errors. In this way, bank stakeholders could come to pay for having trusted supervisors’ positive assessments. It is evident that this approach gives no comfort to advocates of the *laissez faire* approach.

More generally, regulating institutions rather than functions necessarily leads to regulation arbitrage. Increasing regulatory requirements may favour the shift of activities outside the banking sector, with the eventually result that Basel eventually becomes irrelevant.

Finally it is necessary to confront the problem of financial institutions that are too systemic to fail, a problem that has often resurfaced in the previous discussion.

The basic tenet of capitalism is that the best way to impose market discipline is to leave firms free to fail and for their assets to be resolved through bankruptcy. If the orthodox approach favours self-regulation, it should also push for an active action by the state to make bankruptcies socially and politically acceptable. From this point of view, having left the matter to competition policies that apply the notion of contestable markets has obscured the potential systemic effects of bankruptcies. According to that notion, absent barriers to entry and exit, the presence of one or few producers is coherent with a competitive market. In the banking sector this principle has sometimes been mitigated by fixing a maximum market share (e.g. for deposits). This has not avoided the emergence of systemic financial institutions. Just to recall an illustrious precedent of the orthodox approach, the Knight and Simons of the Old Chicago School were very firm in distinguishing between competition and unleashed freedom, the latter giving way to ‘monopolies’ and
dangerous political distortions. Not dazzled by economies of scale and scope, they firmly proposed to force, for all sectors, the ‘one plant’ local firm dimension.\(^{20}\) For the banking sector this should mean allowing only small banks serving local communities which would solve at the same time the problem of size and cross-border banking. Although one may object to this radical outcome, it has the merit of clearly pinpointing the problems and the solution.

As noted above, the current approach favours requiring systemic institutions to have additional capital that should increase their resilience and minimise the negative externalities coming from their resolution.\(^{21}\) A different calibration of capital buffers may come from directly targeting size. For this purpose, capitalisation should be an increasing function of size, at least above a certain threshold, leaving managers and markets to weigh augmented regulatory costs with economies of scale and scope and then decide on the optimal dimension.\(^{22}\) Building on the old prices versus quantities debate (Weitzman, 1974), acting through the market is considered a superior solution with respect to imposing quantitative ceilings. It is held that regulators and supervisors cannot have a general benchmark about the level at which size become systemic, which is not unique and depends on many business and geographical variables. One difficulty associated with this position is how regulators are to calibrate the extra capital buffer if they have no idea of the appropriate systemic size.\(^{23}\) In any case, whether targeting size or resolution, the same problems encountered in the appropriate methodology to determine the level of capitalisation arise.

---

\(^{20}\) See for example Simons 1948, Tonveronachi 1983 and Tonveronachi 1990.

\(^{21}\) The Basel III proposal to tax up to a maximum 2.5% of extra capital buffer the XXL G-SIBs comes from this approach.

\(^{22}\) A bank is defined systemic for both size and complexity. Although in principle complexity is not necessarily linked to size, in practice the mutual relation is significant. Anyway, complexity may be dealt with structural measures such as ring fencing or the Volcker rule. Given that many G-SIBs are predominantly commercial banks, these structural measures do not necessarily solve the problem of size.

\(^{23}\) We are asked to accept a trial and error process also for the effect of extra capitalisation on size. Cfr. Stein 2013.
Together with low confidence on the ex-ante effectiveness of Basel, the aversion to tackle directly the problem of size explains why the focus of regulators is increasingly on resolution. If regulation could make the resolution of Systemically Important Financial Institutions (SIFIs) a swift process without the support of taxpayers, it would be ‘as if’ the fundamental laws of capitalism on bankruptcy were working without disrupting the system. Despite requiring more capital, living wills, extra unsecured debt to be held at the parent or holding level and bail-in instruments, the design of receivership remains, to be optimistic, problematic for cross-border banking and open to loss sharing uncertainty at the national level. The move by the Fed towards US capitalisation and registration of foreign banks, and the debate inside official circles on excessive discretionary powers of the FDIC as the new Orderly Liquidation Authority, show how formidable are the obstacles against orderly and predicable bankruptcy procedures for SIFIs.

To conclude, the major departure of current regulation from what mainstream principles should require concerns the failure to adopt the functionalist approach. Perhaps, as in the USA, this is the effect of having abandoned a structural approach based on institutional specialisation without considering what type of regulation was consistent with leaving financial structures free to adapt; or, as in Continental Europe, the legacy of bank-based systems. In any case, this has increased the regulatory failures coming from a partial analysis approach.

The experience of financial globalisation, a tenet of mainstream economics, has shown the inconsistency between international financial liberalisation and national or regional regulations. In principle, as in the Low Middle Ages, financial globalisation in the context of city-states requires self-regulation and bankers ruling the political roost. The enormous increase of financialisation, also brought about by increased wealth, together with democratic processes in nation-states made bank failures a major national concern. Absent one uniform Empire, political cross-border co-operation is by definition limited by the convergence of national interests. After all, the Ricardian benefits from international trade came from specialisation, not with Portugal mixing with standards on British textile
production, nor with Britain dictating standards on Portuguese wine production. Soft international laws, as international regulatory standards are, are bound to remain effective only if it sets principles and minimum requirements. When experience shows that these requirements are ineffective, the attempts such as those now promoted by G20, FSB, IMF, etc., to strengthen international standards and transform them into maximum convergence are destined to finally clash with structural differences in national interests. As far as the costs remain national, there are good reasons why regulation remains substantially national. With the creation of an international resolution authority unlikely, the current so-called re-nationalisation of international banking is simply the recognition of the limits to international co-operation. The recent European experience of re-nationalisation comes from the absence of a central resolution authority and of a common public backstop. After all, there is yet no convincing evidence of the social gains coming from setting international banks free from locally enforced diversified national requirements. The basic objective should be an international banking system that provides financing for the productive activities of the economy, rather than full convergence on uniform regulations.

These considerations lead to the necessity of a radical rethinking of the regulatory approach.

10. Rethinking the financial structure

One of the foundations of the de-regulation of the last decades was to leave markets free to mould the structure of the financial system through innovations in products and institutions. More precisely, prudential regulation was intended to produce a set of incentives capable of keeping this dynamic process within an acceptable degree of stability, while not interfering with the gains in efficiency coming from competitive innovation. In the aftermath of the recent crisis the G20 decided to guide the re-regulation process; its guiding principle was that financial reforms should avoid ‘excesses’, but not impair private innovations.

As the analyses of the recent crisis and of the ineffectiveness of existing regulation unfolded, the difficulty of designing, within the given regulatory perimeter, the appropriate
system of incentives to constrain unintended consequences within the range of intended results became even clearer. Just to name two of the more evident problems, the rapid development of the shadow banking sector and the shift of banks’ activities from the banking book to securitisation and trading. As discussed previously, mixing the laissez faire approach to the determination of the financial structure with an institutions oriented prudential regulation based on weak foundations leaves such wide margins outside the regulatory perimeter, and to supervisors’ discretion, that the systemic relevance of the unintended consequences easily surpasses that of the intended ones.

The first reaction was to redesign the existing micro-prudential approach in order to strengthen and realign some incentives. Examples are Basel III - with its enhanced requirements for the quality and quantity of capital, its redesign of risk weighting and the introduction of liquidity rules, new recovery and resolution procedures and new rules on OTCs, CCPs and alternative investment funds. However, the nature of the crisis highlighted that the main regulatory failure concerned its disregard of systemic phenomena, which has prompted attention, proposals and legislation on systemic regulation and supervision.

The term systemic includes two dimensions, one related to macro-prudential policy, the other to structural regulation.

In a joint document, the FSB, IMF and BIS (2011) define macro-prudential policy, which complements micro-prudential and macroeconomic policies,

“as a policy that uses primarily prudential tools to limit systemic or system-wide financial risk, thereby limiting the incidence of disruptions in the provision of key financial services that can have serious consequences for the real economy, by dampening the build-up of financial imbalances and building defences that contain the speed and sharpness of subsequent downswings and their effects on the economy; identifying and addressing common exposures, risk concentrations, linkages and interdependencies that are sources of contagion and spill-over risks that may jeopardise the functioning of the system as a whole.” (ibid p. 2)
The definition of the perimeter and tools of macro-prudential policy, and its interactions with micro-prudential and macroeconomic policies, is still a work in progress. For the moment, it has led many jurisdictions to set up an additional regulatory authority, adding complexity to coordination and consistency. The positive aspect is that it has shifted attention from partial to more general analyses, inducing efforts to introduce the financial sector into macroeconomic models; something which presents a revolution for mainstream theory. Reviewing these efforts, Borio (2011) sees the way paved with a number of analytical challenges. “The overriding one […] is to reconsider the prevailing paradigm embedded in macroeconomics. That paradigm is very ill-suited to capture financial instability in a meaningful way and, by implication, also the deeper forces behind all business fluctuations. […] Theoretically, this ultimately calls for a rediscovery of the monetary nature of our economies, in which (inside) credit creation, and hence the creation of purchasing power, plays a key role.” [p. 24]. As will be seen in the next section, downplaying Minsky’s contribution to financial fragility and instability, as Borio himself does, is not the best way to approach that challenge.

In any event, the acceptance of market determined financial structures implicit in the micro-prudential sister of macro-prudential policy is not questioned as a matter of general consistency with other regulatory policies. This means that the transition to structural forms of financial regulation requires recognition of the insufficiency of the micro and macro prudential reforms.

There are two alternatives. One is to rethink anew the entire policy design; the other, the one actually pursued, is to add structural measures to the existing or enhanced prudential framework. Following the latter alternative, structural measures aim to remove substantial residual divergences between the private interests of banks and the public interest. [Vickers 2013, p. 14] Theoretically, we are in the realm of the second best approach, introducing further imperfections with respect to the ideal model but requiring as always cost-benefit evaluations.
At present, discussions revolve around three schemes: the Volcker rule, Liikanen subsidiarisation and Vickers ring fencing.\textsuperscript{24} Their lowest common denominator is the protection of taxpayers from shouldering the costs of banking crises, in particular by making private bank resolution possible.

The Volcker rule forbids banks from engaging in proprietary trading and limits participation in hedge and equity fund activity. The result is complete institutional separation, with the excluded activities managed by intermediaries that cannot acquire insured deposits and cannot be part of a bank holding company. The idea is not to purge banks from their riskiest activities (securities underwriting and market-making are maintained), but to exclude risky speculative activities that do not serve the public purpose.

The Liikanen subsidiarisation is wider in scope than the Volcker rule, but less strict in separation. Proprietary trading and market-making above certain size thresholds have to be placed in a separate legal entity (Investment Bank IB), which is barred from being funded by retail deposits. Both types of entities (Commercial Bank CB and IB) may be housed in the same holding company structure, but each must comply with regulatory requirements on a stand-alone basis. Intragroup exposures would be subject to the same regulatory limits as when dealing with external counterparties. CBs would be permitted to engage in some investment bank business, such as underwriting and wholesale activities – although with limits on wholesale funding –, and international cross border activity. Supervisors could shift the separation line of permitted activities if deemed necessary for resolution purposes. The separation of activities is more clear-cut than in the Volcker rule, but with prudential regulatory walls as substitute for complete institutional separation. The explicit goal is to preserve the economies of scope across products of the universal banking model, while, as for the Volcker rule, trying to protect depositors and taxpayers from unnecessary risky activities.

\textsuperscript{24} The following analysis takes as reference the original schemes and not the modifications that are being introduced in the process of converting them into actual legislation.
The Volcker rule and the Liikanen subsidiarisation have limited effects on bank size and complexity, hence on rendering resolution a swift process.

The Vickers ring fencing shares with the Liikanen subsidiarisation the proposal for prudential regulatory walls, but pursues goals that are more stringent. The necessity to smooth resolution processes and shield taxpayer’s financial exposure to crisis bailouts is coupled by ring fencing the payment system and loans to households and Small and Medium Enterprises (SMEs). Thus, ring-fenced banks would limit their activities to retail banking, for which they would have a sort of monopoly, and would be barred from international banking, an activity only permitted to non-ring fenced banks. The result is separation from non-retail entities and from international crises. Foreign intermediaries working in the host retail market would be subject to ring fencing requirements. This would permit application to all banks working in a given jurisdiction of stricter regulatory requirements without fear of reducing their competitive edge. The scheme has several merits. It deals more forcefully than the other schemes with complexity and size, although only for ring-fenced banks. It avoids cross-border resolution problems for ring-fenced banks. If the regulatory prudential walls are effective, it shields retail banking from bankruptcies of non-ring fenced entities pertaining to the same group. If the commitment not to use public safety nets for non-ring fenced entities is credible (a big if when they are systemic), the scheme simplifies resolution procedures and limits the eventual recourse to public funds to basic financial needs of the home market. Its more relevant limit is that non-ring fenced entities may have a relevant role in short-term financing of the economy, may be systemically relevant and therefore moral hazard problems related to implicit public guarantees would not disappear.

A full evaluation of the effects of the previous schemes cannot overlook the more general regulatory context in which they are thought to be applied. For example, if we consider the

---

25 The rule also forbids to ring-fenced banks to own or participate in subsidiaries and branches in foreign markets. As long as the UK remains a member of the EU, or the proposal were adopted at the EU level, international banking means outside Europe.
Volcker rule together with United States Federal Reserve Regulation W, which already limits intragroup exposures, and the Fed proposal on the US establishment of foreign banks, we have a mix of separation and ring fencing with more relevant consequences on bank resolution than looking solely at the Volcker rule. Besides, the Financial Stability Oversight Council FSOC, in difference from the European Systemic Risk Board ESRB, may force additional structural measures on systemic intermediaries. However, this relates to the discretionary powers of supervisors, not to rules.

The three schemes have encountered strong opposition from the banking industry, mainly on grounds of producing diseconomies of scope and higher funding costs, unnecessarily so when considered together with the array of the prudential reforms that are being implemented. Their approach is to ask the implementation be delayed until the results of the prudential reforms can be evaluated. The difficulties associated with this type of discussion concerns how to measure private costs and social benefits and where to put the balance between them.\textsuperscript{26} As the recent experience in the USA and EU shows once again, this approach necessarily leads to political compromises being dominant over fundamental principles in the determination of final decisions.

11. A heterodox approach to financial fragility and regulation

As noted above, one of the basic drawbacks in the current approach to financial regulation is the absence of any clear connection between theoretical foundations explaining the operation of the economic system and regulatory proposals to ensure that the operation of that system is financially stable. This is perhaps not surprising, for as noted above and as Hyman Minsky has pointed out in his Financial Instability Hypothesis, traditional equilibrium theory has no place for financial crisis as a natural outcome of the operation of the system. Excluded by definition of the self-adjusting equilibrium system, financial crisis can only be the result of external factors or shocks, or malfeasance and which are dealt with by means of ad hoc, backward-looking, casuistic regulations. Yet, the modern

\textsuperscript{26} Discussing the topic of structural reforms, Gambacorta and van Rixtel (2013) have recently reviewed researches on economies of scale and scope in banking, showing the inconclusiveness of their results.
financial system is driven by profit seeking institutions, ever evolving new operating mechanisms through financial innovation to gain competitive advantage, either relative to other institutions, or by reducing the constraints on activity of existing regulations. Thus, a process of Schumpeterian competition and regulatory arbitrage will be simultaneously and continuously changing the structural character of the system and its mode of operation. The implication of this alternative approach is that regulation must then also be a dynamic, evolving institutionally grounded process.

Minsky sought to remedy the absence of a theoretical foundation for financial regulation through his approach to Financial Instability, extending Keynes’s analysis of income and expenditure generation in the General Theory to provide an explanation of the cyclical evolution of the financial system. Instead of starting from individual preferences and endowments to seek the vector of prices that achieves the optimal Pareto equilibrium allocation of given resources through free market exchange, Minsky starts from the basic characteristic of modern capitalism that the control of capital assets is achieved via the issue of liabilities. The balance sheets of the major actors of the private sector of the economic system are thus the building blocks of economic analysis.

Since liabilities require cash payments and assets produce cash inflows, the viability of any balance sheet structure will depend on the balance of cash inflows and outflows. If inflows are always greater than outflows, then the assets being financed have positive net present value and the value of assets will exceed the value of liabilities. This is the basic condition for the stability in the economic system, positive net present value for investment projects and positive net worth for the balance sheet, or a positive income statement and balance sheet. In other terms, the income from the financed assets should always be capable of covering the payments on the liabilities held by the lenders. And since lenders will hold these liabilities as assets that are in their turn financed by liabilities, each sector of the economy will be linked through an interlocking system of receipts and payments. Since each sector’s liabilities will usually appear as assets on the balance sheet of some other
sectors this approach highlights the interconnectedness, or complexity, of payment commitments as a crucial element in assessing the financing structure.

The financially stable, successful operation of any financial or non-financial institution thus presumes the ability to issue liabilities to acquire productive assets and that those assets generate incomes that are sufficient to cover the service on the debt that finances them. This provides the interface between business and finance, and the characteristic feature of banks in the financial system.

With respect to this basic characteristic of capitalist production, Minsky points out that any actor can issue a liability, an IOU, that promises future payment in exchange for present purchase, but that not all agents are able to get them accepted in practice for goods and services. Business firms in general cannot do this, nor can private individuals. But it is the characteristic of the business of banks to issue their own liabilities, promises to pay, that serve as generally accepted means of payment. Thus, Minsky notes that, “the fundamental banking activity is accepting, that is, guaranteeing that some party is creditworthy,... A bank loan is equivalent to a bank buying a note that it has accepted.” (1986, 256) The business of banking is thus the “acceptance” and validation of business liabilities which it holds as assets on its balance sheet in exchange for granting its own deposit liabilities to the borrower. The bank is able to charge interest on this transaction because the borrower exchanges an illiquid liability for a liquid liability that can be used as means of payment. That these deposits can function as a means of payment provides the borrowing firm with the ability to acquire the capital assets it needs despite the inability to purchase them with its own liabilities. Capital investment thus depends on the ability of the banking system to issue liabilities that serve as generally acceptable means of payment. Indeed, any business activity, including the acquisition of inputs and the payment of labour will require the intermediation of the banking system to provide the necessary means of payment. Thus, the primary function of banks is not intermediaries between household savers and business investors, but as the essential creators of liquidity that allows the capitalist process of
production and investment which relies on expenditure today producing income in the future.

This approach is a radical departure from traditional neoclassical theory. Rather than explaining the existence of banks as being due to their efficiency in acting as a conduit between savers and investors because they have better information concerning investments than private savers, banks are seen as providing to business firms the ability to acquire the assets they need to engage in production or investment in new productive capacity and which will eventually produce the income from which savings will arise. Banks convert illiquid liabilities into liquid means of payment and receive income for this function. The investment of household savings is the responsibility of other financial institutions and advisers, not banks.

Thus, the importance of the banking system is its ability to allow firms to issue liabilities that can be used to acquire productive assets, whether labour and inputs, or capital equipment. But how are banks able to do this? In Minsky’s words (1988, 258), “the only reason why banks are special is that they operate the “ultimate” payment system within economies .... There are now alternatives to banks for all but the provision of the ultimate payment mechanism function. Because banks operate the ultimate payments mechanism, those liabilities of banks which serve as the “medium of exchange” also serve as the standard in which domestic public and private debts are denominated.” Thus, if banks did not operate the payments mechanism providing clearing for their deposit liabilities, they could not convert business liabilities into means of payment. This function is often called intermediation, the conversion of short-maturity bank liabilities into longer maturity business liabilities. But this is to miss the basic point of intermediation which is to convert illiquid business liabilities into perfectly liquid means of payment: sight deposits transferable by check. And the income that banks earn in the form of net interest margins is the result of the difference between the liquidity premium on the liabilities it issues and those it acquires from borrowers. Thus, there is also a profit incentive for banks to keep their liabilities as liquid as possible.
This characterisation of banking operations highlights banks’ two basic functions: they are liquidity machines vouching for the creditworthiness of the issuers of private liabilities, and they do so because they operate a generally acceptable payments system for their liabilities. Banks are important because their balance sheets hold the liabilities of the productive sector that are issued to acquire control over the productive assets of the private business sector, and they issue deposit liabilities that are held by the general public because they can be used as final means of payment. Thus, the balance sheet of the financial sector represents a combination of the financing of the capital development of the economy, as well as the means of payment and savings of the general public.

As a result, Minsky characterises the role of regulation in the financial system as having to serve two, mutually conflicted, masters: “any capitalist banking and financing system” is “drawn between two masters” that it “needs to serve: one master requires assurance that the financing needed for the capital development of the economy will be forthcoming and the second master requires assurance that a safe and secure payments mechanism will be provided.” “It ... needs to be understood now that development financing involves taking risks ... The need is for a regulatory and supervising authority for the financial system that accepts that financing development opens the system to losses that have the potential for adversely affecting the safety and security of the economy’s payment facilities. To allow for this possibility the regulators need to try to insulate the payments system from the consequences of such losses. The problem therefore is to provide for protection of the payments system from the consequences of the losses which may ensue from development financing.” [Minsky, 1994, 10-11].

From this point of view, the fundamental role of prudential banking regulation is to ensure that banks are able to provide this dual function, but it is also inherently contradictory. Banks that never made losses would not be providing the support of entrepreneurial innovation that is the driving force of capitalism. For example, the government of virtually every nation reserves for itself the monopoly right to create means of payment in the form of coin and currency. As means of payment, aside from the risk of theft or loss, it is
perfectly safe and secure. It can be used by government to make purchases and as such represents creation of liquidity for the government itself, but it cannot create liquidity for the liabilities of private sector investors unless it is organized as a government-owned bank. And governments have done this in the past, and many still own and operate central banks or private development banks, but most rely on the private banking system to provide for the financing of risky investments. Thus, government liabilities serve the payments function well, but not the creation of liquidity for private business seeking to finance of productive investment.

Thus, the banking system must compete with the government liabilities, notes and coin, as means of payment and the government liabilities serve as the units of account for the private payments system. As a result, part of the operation and control of the bank payment system is for banks to provide a guarantee that bank deposits will always be convertible into government liabilities on a one to one basis. To do this, banks will also create deposits against the collateral of government liabilities. This means that individual banks will create two distinct types of deposit liabilities, one which is fully collateralized by government liabilities and the other which is collateralised by the private liabilities of a business borrower and subject to risk of future repayment. In this way the dual function is also reflected in the creation of deposits as private means of payment.

Thus, the ability of private banks to provide this dual function depends on their ability to maintain their monopoly over the payments system. As Minsky noted, “In our system payments banks make for customers become deposits, usually at some other bank. If the payments for a customer were made because of a loan agreement, the customer now owes the bank money; he now has to operate in the economy or in financial markets so that he is able to fulfil his obligations to the bank at the due dates. Demand deposits have exchange value because a multitude of debtors to banks have outstanding debts that call for the payment of demand deposits to banks. These debtors will work and sell goods or financial instruments to get demand deposits. The exchange value of deposits is determined by the demands of debtors for deposits needed to fulfil their commitments. Bank loans, while
ostensibly money-today for money-later contracts, are really an exchange of debits from a bank’s books today for credits to a bank’s books later.”

The stability of the banking system thus depends on the ability to meet debt service and to repay outstanding loans. There are several types of risk to this stability.

• For the individual bank, deposit withdrawals may exceed repayments of loans. This is known as deposit drain and threatens the first function of the bank in providing a safe and secure payments system and its ability to maintain perfect substitution with government liabilities.

• For the individual bank, the debits on the bank’s books today may fall short of the credit to the bank’s books later. This is the problem of non-performing loans, which is the second function provided by banks in financing loss-making investments.

• For banks as a whole, all the deposits created by all the banks must be on the balance sheet of some bank in the system, so the first risk noted for individual banks does not exist unless deposits are converted into government liabilities of coin and currency. This is known as a systemic liquidity crisis and represents a breakdown of the banks’ control over the payment system.

• And if the future credits of the system as a whole do not meet the current debits, it may create financial crisis and systemic insolvency and represents a breakdown of the risk-taking function.

Prudential measures thus seek particular regulations to manage these threats to financial stability. For the first risk of deposit drain, banks either have voluntarily or are required to hold reserves of vault cash, deposits with the central bank which can be converted into government cash at will, or offered access to the discount window of the central bank to sell their performing loans for government liabilities. For the second type of risk to stability, banks will be required to hold capital sufficient to absorb the losses on bad loans or be wound up if losses are in excess of capital. For the third type of risk, the banks may be offered access to borrowing government currency at the central bank. For the fourth
type, if bank capital is exhausted by the losses government must decide to nationalize the banks or to recapitalize them.

Banks will seek to economise on reserves in excess of the legal or business requirements by holding secondary reserves of assets that can be quickly converted into government cash or in the case of a legal requirement arranging the balance sheet or creating types of instruments that are not covered by the reserve requirement. But the ability of the former measures to provide liquidity depends on market conditions, in particular the ability to find a buyer of the assets or someone willing to lend cash against them. In conditions of liquidity shortage by an individual bank, counterparties may be unwilling to transact. In case of generalised illiquidity all institutions will be seeking to convert their secondary assets into cash, which means that none of them can except by using them as collateral against a loan from the central bank. Secondary reserves are thus only liquid in normal conditions and do not in general provide assurance of being liquid in crisis. In the case of the latter, they impede the ability of generate liquidity when it is needed and increase the risks of instability due to illiquidity.

There are alternative means of providing these types of regulation. The Robertsonian aphorism that all the money that is anywhere must be somewhere leads to the realisation that the banking system as a whole cannot be short cash as long as a deposit that is redeemed at one bank is re-deposited at another bank in the system. A run against a single bank is then just a misdistribution of reserves in the system. Thus, in many countries banks have sought mechanisms to pool reserves so that they can be transferred from banks short of reserves to banks that are long reserves. More evolved examples are bank Clearing Houses that have been major mechanisms of preserving the banks’ control over the payments system. They provide an example of self-regulation by means of the conditions of solvency and inspection that they place on membership. Some Clearing Houses have provided for the creation of internal Clearing House credits to meet specific short positions in times of liquidity crisis, providing endogenous liquidity that acted as a private lender of last resort system. This was the system in place in the United States before the creation of
the Federal Reserve. Indeed, “reserve” appears in the title of the US central bank because the initial proposals sought to create an institution that would pool bank reserves and provide a mechanism to respond more effectively to a run on an individual bank in the system.

Such institutions preserve the payments system when single banks face difficulty in preserving their commitment to perfect substitution of their own liabilities with government currency. However, when there is a severe liquidity crisis that manifests in the form of a run on the entire banking system and the public only wants to hold government liabilities, such private mechanisms are inadequate and either direct government intervention or indirect liquidity creation via a central bank discount mechanism is required. This is generally characterised not as a “bank run,” but as a “run to currency,” which only the provider of government guarantees to payment can control.

Thus, ultimately the ability to guarantee the perfect substitution of deposits and cash depends on a financial institution’s ability to access government liabilities, either directly from the government or indirectly through the swap of bank assets for government liabilities on the balance sheet of another banks or the central bank. This is another example of why banking is always a public-private partnership.

An alternative mechanism of ensuring the integrity of the payment system is deposit insurance which is primarily designed to supplement bank capital and preserve the value of deposits in the payment function in the case of individual bank insolvency. In a sense, it seeks to separate the role of bank capital and the reserves in preventing instability. The logic is that if a bank is insolvent, because its expected future credits fail to materialize to cover existing debits by more than its capital, that is that it makes bad loans, secured by its own credit, it is the borrower’s initial deposit created to fund the loan that should be responsible for the loss. Then, the loss should be borne by the bankers who made the loan and the shareholder-creditors to the bank. Thus, in the case of losses produced by the failure of business borrowers to meet their commitment, losses should be met by the business borrower first, and then by the bank shareholders represented by the bank’s
capital. This is the basis of the arguments against public “bailouts” of banks and the insistence that the costs of bankruptcy due to bad loans should be met first by the bad borrower, and then by the bad banker’s shareholders; if this does not cover losses other creditors of the bank should be included. But, in order to preserve the payments system and avoid contagion in the form of a bank run on other banks, the bank’s depositors should be protected in order to ensure a safe and sound payments system.

The key to the acceptability of deposits in the banking system is thus that borrowers must be in a position to make timely repayment to the lending bank by transferring a deposit. Thus, the deposit the borrower receives from the bank is only useful if it can be used to acquire the inputs required in production. Used as a means of payment, the ownership of deposits is transferred to households who are free to deposit them in any other bank in the system. This is what is called “deposit drain”, and generates the need for bank reserves which are used to effectuate the transfer of ownership of the deposit to another bank. This is achieved via a transfer of reserves in the interbank market. The deposit originally created by a loan thus appears at the acquirer’s bank as a transfer of reserves. There is thus no way of distinguishing it from an original deposit of notes and coin. It is thus also the case that there is no way of distinguishing between an original deposit of reserves and a deposit created with a bad loan or a current loan.

In a modern financial system it is even more difficult to make these distinctions between types of deposits. The main reason is that generally the way government currency enters the financial system is not through the sale of goods and services to government in exchange for cash. Most cash in circulation results from the conversion of a deposit, created by a bank loan, into cash by a bank converting its own reserves into cash or using its vault cash. And most government expenditures are paid via the transfer of a deposit at the central bank to the sellers’ deposit account. As such, the idea of a deposit resulting from the direct deposit of currency becomes moot, and defeats the idea that setting a limit on deposit insurance eliminates the moral hazard created by insurance.
But there is a further difficulty created by deposit insurance and that is that it has normally been provided as insurance funded by the payments of a premium by the insured. Some are public funds, as in the US, some are private as is the case in some European countries. As long as the insurance was meant to counter the contagion that might result from the difficulties facing a single institution the insurance principle was viable. But, in the case of systemic risks to the entire banking system because banks have achieved a very large size, the insurance principle is no longer credible. Minsky argues that to preserve the integrity of the system it must take on the nature of a government guarantee for all deposits in the banking system. Indeed, this is what was required in the recent European experience and which is a point of contention in the formulation of a common EU insurance fund which to retain credibility would require a government guarantee. But since there is no EU government capable of such action it would implicitly falls to the ECB to provide such a guarantee.

The need to meet the promise to transform deposits into coin and currency for the creation of credit highlights the importance of reserves. The deposit created by a loan could not be transferred to another bank and thus could not be used as a means of payment were it not for the transfer of reserves between banks. And, as Minsky points out, an equally important part of the ability of banks to create credit is that after the borrower has used the deposit as a means of payment, it has to acquire deposits created by other banks [or its own] in order to extinguish the loan. Thus, there is a demand for deposits from the banks both to acquire reserves and to repay loans, while the household demand for deposits is independent of their origin.

As noted above, deposits have formed the basis of the regulatory system since the end of the 19th century. But since banking is a private sector profit based system it has been subject to Schumpeterian innovation and competition. Just as deposits replaced the issue of private bank notes, other types of payment system have evolved as financial institutions have sought inroads into commercial banks’ monopoly of insured deposits under Glass-Steagall Act. Indeed, the history of financial deregulation in the United States has been a
story of such innovation in payments systems from Negotiated Orders of Withdrawal for savings banks to bank credit card systems and the creation of check drafts on stock brokerage accounts. Indeed, many of the difficulties in the US surrounding deregulation have been caused by the attempt to regulate institutions or particular products rather than the function that banks provide in operating the payments system.

This is also true of the other function of banks, the creation of liquidity to finance business sector investments. Indeed, the provision of liquidity via the acquisition of a business liability through the creation of a deposit liability has long since ceased to be the primarily source of liquidity creation in the economic system. Indeed, competition in new innovative forms of liquidity creation has been presented as the evolution of the “shadow banking” system. However, this denomination is only relevant because the traditional model for regulation has been of deposits issued by commercial banks to business borrowers. Indeed, much of the instability that was experienced in the recent crisis might have been avoided, if regulations had focused on the function of financial institutions in producing innovations in the creation liquidity in competition with and in substitution for commercial bank deposits.

In a competitive, profit driven financial system institutions will be innovating in the provision of payments systems and in the provision of liquidity, the dual function of any financial system. A regulatory system that focusses on the particular products or institutions that provide these functions will always be fighting the last war and failing to identify sources of fragility in the system. Indeed, increasingly these functions are no longer provided within the same institutional structure.

In addition to the provision of liquidity, financial institutions also provide for the intermediation of savings and investors, but this takes place through capital markets populated by financial institutions that provide a different type of liquidity. The role of capital markets is to provide long-term finance for capital projects and research and development in support of innovation in the financial system. As Keynes reminds us, the volatility of such projects over the longer term is impossible to evaluate for the saver.
seeking a safe return on savings. The capital markets thus provide for a bridge between the long-term financing needs of risk-taking businesses and the preference for more secure short-term financing of the investing public by providing a deep and liquid market in long-term assets that makes them saleable on a short-term basis and thus liquid. Financial institutions do this by acting as market-makers offering to buy and sell at fixed prices up to certain volume limits. They thus use their own capital to provide capital market liquidity, which is inherently different from the liquidity provided by commercial banks. There are many different types of financial institution that provide this liquidity to the market and they should by the same token be regulated according to their function in providing market liquidity.

12. Conclusions

It is interesting to note that the 1930s Glass-Steagall legislation that is widely considered to have been highly successful in providing financial stability over the period from 1940 to 1970 was based on regulations that imposed an identity between functions and institutions. Commercial banks as institutions were limited to the basic function of liquidity creation, while investment banks were limited to capital market liquidity creation and the associated activity of securities trading which was required to provide that function. This regulatory regime broke down when this imposed identity was lost due to deregulation and competition which eventually produced a system in which all institutions could provide all types of financing services in the 1999 Financial Services Modernisation Act. The result was a loss of focus on the evolution of the functions of liquidity creation and thus on the creation of instability in the system which eventually produce the 2007-8 crisis and the Great Recession.

This is not to argue that financial institutions cannot operate across a wide range of financial activities. German universal banks were highly stable institutions while providing the full range of financial services. But this was primarily due to the fact that German financial regulations differentiated the different functions of a bank’s balance sheet and regulated them accordingly. The regulation was effectively by function within the universal
bank. Thus, the conclusion that concentration on functional regulation should allow for the proliferation of different types of financial products and financial organisations as long as they are regulated according to their function in providing liquidity of different types to the financial system.

13. References


Financialisation, Economy, Society and Sustainable Development (FESSUD) is a 10 million euro project largely funded by a near 8 million euro grant from the European Commission under Framework Programme 7 [contract number : 266800]. The University of Leeds is the lead co-ordinator for the research project with a budget of over 2 million euros.

THE ABSTRACT OF THE PROJECT IS:

The research programme will integrate diverse levels, methods and disciplinary traditions with the aim of developing a comprehensive policy agenda for changing the role of the financial system to help achieve a future which is sustainable in environmental, social and economic terms. The programme involves an integrated and balanced consortium involving partners from 14 countries that has unsurpassed experience of deploying diverse perspectives both within economics and across disciplines inclusive of economics. The programme is distinctively pluralistic, and aims to forge alliances across the social sciences, so as to understand how finance can better serve economic, social and environmental needs. The central issues addressed are the ways in which the growth and performance of economies in the last 30 years have been dependent on the characteristics of the processes of financialisation; how has financialisation impacted on the achievement of specific economic, social, and environmental objectives?; the nature of the relationship between financialisation and the sustainability of the financial system, economic development and the environment?; the lessons to be drawn from the crisis about the nature and impacts of financialisation?; what are the requisites of a financial system able to support a process of sustainable development, broadly conceived?’
THE PARTNERS IN THE CONSORTIUM ARE:

<table>
<thead>
<tr>
<th>Participant Number</th>
<th>Participant organisation name</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 [Coordinator]</td>
<td>University of Leeds</td>
<td>UK</td>
</tr>
<tr>
<td>2</td>
<td>University of Siena</td>
<td>Italy</td>
</tr>
<tr>
<td>3</td>
<td>School of Oriental and African Studies</td>
<td>UK</td>
</tr>
<tr>
<td>4</td>
<td>Fondation Nationale des Sciences Politiques</td>
<td>France</td>
</tr>
<tr>
<td>5</td>
<td>Pour la Solidarité, Brussels</td>
<td>Belgium</td>
</tr>
<tr>
<td>6</td>
<td>Poznan University of Economics</td>
<td>Poland</td>
</tr>
<tr>
<td>7</td>
<td>Tallinn University of Technology</td>
<td>Estonia</td>
</tr>
<tr>
<td>8</td>
<td>Berlin School of Economics and Law</td>
<td>Germany</td>
</tr>
<tr>
<td>9</td>
<td>Centre for Social Studies, University of Coimbra</td>
<td>Portugal</td>
</tr>
<tr>
<td>10</td>
<td>University of Pannonia, Veszprem</td>
<td>Hungary</td>
</tr>
<tr>
<td>11</td>
<td>National and Kapodistrian University of Athens</td>
<td>Greece</td>
</tr>
<tr>
<td>12</td>
<td>Middle East Technical University, Ankara</td>
<td>Turkey</td>
</tr>
<tr>
<td>13</td>
<td>Lund University</td>
<td>Sweden</td>
</tr>
<tr>
<td>14</td>
<td>University of Witwatersrand</td>
<td>South Africa</td>
</tr>
<tr>
<td>15</td>
<td>University of the Basque Country, Bilbao</td>
<td>Spain</td>
</tr>
</tbody>
</table>

The views expressed during the execution of the FESSUD project, in whatever form and or by whatever medium, are the sole responsibility of the authors. The European Union is not liable for any use that may be made of the information contained therein.

Published in Leeds, U.K. on behalf of the FESSUD project.