THE | C. BOYDEN | GRAY | CENTER

Taking Systemic Risk Seriously in Financial Regulation

M. Todd Henderson
James C. Spindler

CSAS Working Paper 16-13
Financial Regulation: Political, Administrative, and Constitutional Accountability
December 9, 2016

Antonin Scalia Law School
C. BOYDEN GRAY CENTER FOR THE STUDY OF THE ADMINISTRATIVE STATE
George Mason University

Taking Systemic Risk Seriously in Financial Regulation

M. Todd Henderson & James C. Spindler[·]

· Todd Henderson is Professor of Law and Aaron Director Teaching Scholar, The University of Chicago Law School; James Spindler is The Sylvan Lang Professor, University of Texas Law School and Professor, University of Texas McCombs School of Business, respectively.

Taking Systemic Risk Seriously in Financial Regulation

Abstract

Bank regulation failed in the run up to the Financial Crisis of 2008, as it has numerous times in the course of US history. This is despite the existence of traditional prudential regulation (such as capital adequacy mandates, reserve requirements, and bank examination) as well as more common legal remedies, such as tort and contract litigation. Unsurprisingly, in the wake of these failures, many reforms have been proposed, and some adopted, to try to reduce bank risk-taking. These reforms include limiting bank size, requiring bank managers to be paid differently, restricting investment in high risk financial products, and, of course, tightening up existing prudential regulation.

In this paper, we first categorize these proposals into traditional categories of regulation—ex ante and ex post forms—and point out the weaknesses of each. Ex post regulation – generally, liability after the fact for harm caused – fails almost by construction: given externalities of systemic risk and leverage, judgment-proofness is virtually guaranteed and is uninsurable. Ex ante regulation – which comprises the bulk of current prudential relation – is, as a starting point, inefficient because it fails to take into account both private information and subsequent public information. More vexingly, ex ante regulation encourages worse behavior: size limits and transactions taxes encourage higher-octane bets, and asset restrictions lead to the recreation of the same risk profiles in less efficient ways.

We then describe an intermediate form, what we call the "regulatory veto," which allows regulators to intervene to reduce bank risk taking after banks have started their activities but before the losses have occurred. We show how the regulatory veto is, potentially, an elegant solution to the information problem presented by ex ante regulation and the judgment-proofness problem of ex post regulation of bank activities. However, the regulatory veto is subject to a structural flaw: banks get to move first in a form of "Ultimatum Game" and choose supra-optimal levels of bank activities, which are not quite bad enough to cause regulators to shut them down. To mitigate this flaw, we propose reforms to enhance regulators' ability to credibly commit and to reduce banks' ability to game the system.

Table of Contents

I. INTRODUCTION **4**

II. THE BANKING SYSTEM: A PRIMER **11**

A. BANKING BASICS – LONG-SHORT MISMATCH, RUNS, AND MAINTAINING CONFIDENCE 12
B. BANKING REGULATION BASICS 17
1. Prudential Regulation *19*
2. Capital and Reserve Regulation *22*
3. The Regulatory Veto *24*
C. REGULATORY REFORM PROPOSALS 29

III. A SIMPLE MODEL OF REGULATION OF BANK RISK-TAKING **32**

A. THE MODEL 34
1. The banking firm and its projects *35*
2. Shareholders *35*
3. Bank creditors *36*
4. The regulator and systemic risk *37*
B. HOW DO THE VARIOUS ALTERNATIVES FARE? 39
1. A baseline: laissez-faire *39*
2. Ex post: litigation and insurance *41*
3. Ex ante regulation: size and asset type restrictions *45*
4. A middle ground: the regulatory veto *58*

IV. HOW TO MAKE BANKING REGULATION BETTER **65**

V. CONCLUSION **67**

I. INTRODUCTION

Modern American bank regulation is an array of devices designed to limit risk at individual banks and in the banking system as a whole. Allowing banks to take deposits and other short term liabilities in order to invest in loans and other long term financial assets creates wealth, but private wealth-generating activities may not be socially valuable if the costs they externalize onto others exceed the benefits. Banking regulation, therefore, attempts to attain the socially optimal level of banking activity by limiting the activity of banks.[1] Departures from the optimum hurt society: too little lending lowers bank profits and reduces the economy's access to credit, while too much lending and investment creates excessive risks that will largely, in the event of catastrophic failure, be borne by someone else.

The regulatory tool kit of bank regulation is diverse and includes both what we term "ex ante" and "ex post" forms of regulation. Consider first ex ante bank regulation (often referred to as "prudential" bank regulation"). Entry into the banking business is regulated by state and federal agencies. The kinds of activities banks can engage in—the "business of banking" and activities "necessary" to this business— are limited.[2] "Safety and soundness" rules limit the types and sizes of loans, restrict lending and other dealings to and with bank executives, mandate amounts of cash that must be held in reserve,[3] and prescribe limits on the amount of leverage that banks may take.[4] Disclosure obligations under

---

[1] For welfarists, this is the purpose of regulation of any activity. As discussed below, the case for active and aggressive regulation is stronger in the banking context than in other areas. While it might make sense to let our regulation of most consumer goods be handled through voluntary market transactions and an ex post remedy for fraud or other harm, this approach is unlikely to work for banking.

[2] See The National Bank Act, 12 U.S.C. § 24(7).

[3] Various state and federal regulators, such as the Department of Treasury or the Office of the Comptroller of the Currency, prescribe reserve requirements, cash on hand banks must maintain to repay depositors and creditors. For example, the Federal Reserve Board's Regulation D prescribes reserve requirements beyond those required by law. See http://www.federalreserve.gov/monetarypolicy/reservereq.htm. For example, the Garn-St. Germain Act exempts the first $2 million from reserve obligations, and the first $25 million above this is subject to a relatively low (that is, 3 percent) reserve requirement by the Monetary Control Act of 1980. Regulation D sets the reserve requirement above $89 million at 10 percent. See *id.*

[4] Specifically, capital adequacy requirements, primarily under the Basel Accords, limit the bank's ability to engage in risky activities that generate obligations the

Basel Pillar 3 mandate disclosure to enable "market discipline" of banks.[5] Federal deposit insurance mandates that banks pay premiums to insure demand deposits against deficiencies.[6] Finally, various government agencies have broad statutory power to seek injunctive relief and damages in the event of losses caused by banking. So far, banking regulation looks a lot like other areas of regulation, albeit far more comprehensive and more strict.

Ex post measures generally involve mopping up after the damage has been done. Jilted creditors and counterparties may pursue their remedies in court, as may the government, particularly in the case of a federal payout. Banks that fail are forced into receivership, their assets unwound and creditors paid off to the extent possible. Finally, bank personnel who have committed wrongdoing may be sued or prosecuted.

In addition, there is a rather unique banking regulatory institution that spans ex post and ex ante measures. Specialized regulators, known as "examiners," are assigned to individual banks.[7] Examiners have an additional, and unusual, mandate: to assess the riskiness of a bank's entire bank portfolio of activities and to shut down the whole bank or specific bank activities if the

---

bank may be unable to repay. For a discussion of the Federal Reserve's implementation of the Basel accords, see Joint notice of proposed rulemaking implementing the Basel II standardized risk-based capital framework in the United States, June 26, 2008, Docket number R-1318, available at www.federalreserve.gov/bankinforeg/basel/USImplementation.htm#baseIIa.

[5] *See* Basel II, Pillar 3, at 175 (available at http://www.bis.org/publ/bcbs107c.pdf). For an example of Pillar 3 disclosure, see, e.g., Bank of America Pillar 3 disclosures at http://investor.bankofamerica.com/phoenix.zhtml?c=71595&p=irol-basel#fbid=5xLoxBRdnYn

[6] *See* 12 U.S.C. § 1811, and regulations promulgated thereunder. See also http://www.fdic.gov/deposit/deposits/.

[7] According to the Wamu failure report: "FDIC assigns a dedicated examiner to the largest insured financial institutions. The dedicated examiner serves as the case manager for these institutions and works in cooperation with primary supervisors and bank personnel to obtain real-time access to information about an institution's risk and trends." *See* Dept. of Treasury, Offices of Inspector General, "Evaluation of Federal Regulatory Oversight of Washington Mutual Bank," Report No. EVAL-10-002, Apr. 2010 ("WaMu Report"), at 69. To assess risk at a broader level, FDIC conducts a wide range of activities to monitor and assess risk from a regional and national perspective. At the institutional level, FDIC monitors large non-FDIC supervised institutions primarily through its Dedicated Examiner and Case Manager Programs. FDIC relies on the PFR's examinations to determine a bank's overall condition and the risks posed to the Deposit Insurance Fund. Additionally, FDIC, by statute, has special examination authority and certain enforcement authority for all insured depository institutions for which it is not the PFR. See *id* at 67.

expected costs and harms of the activities are excessive.[8] In theory, bank examiners ought to be able to zero in on the activity sweet spot at which social welfare is maximized. While the nature and scope of examinations differ, for large, complex banks, the examination team is a constant regulatory presence.[9] Regulators have vast powers to curtail bank activities, to require increased monitoring of activities, to require the bank to raise more capital, to impose new operating procedures, to replace management, and to even shut down specific activities or the entire bank.[10]

Notwithstanding the panoply of regulatory measures available, this regulatory system utterly failed in the run up to the Financial Crisis of 2008, as many banks and financial firms collapsed or had to be bailed out as the result of too much exposure to residential real estate, especially derivative instruments linked to subprime loans.[11] This was not the first time the system failed to keep bank activities in check. During the Savings & Loan Crisis of the late 1980s, certain types of banks, known as thrifts, engaged in far too much lending to particular borrowers, resulting in widespread bank failures.[12] In just the two most recent crises, the estimated direct losses exceed half a trillion dollars.[13] The indirect costs and broader social costs are probably many times this amount.[14] Banking crises appear to be a recurring

---

[8] For a detailed discussion of the supervision process, see M. Todd Henderson & Frederick Tung, *Pay for Regulator Performance*, 85 S. CAL. L. REV. 1003 (2012); *see also, e.g.*, "An Examination of the Banking Crises of the 1980s and Early 1990s," ("FDIC Report") available at www.fdic.gov/bank/historical/history/421_476.pdf. *See also* WaMu Report, *supra* note 7, at 66.

[9] For instance, during the period leading up to WaMu's failure, federal bank examiners spent over 160,000 hours (about 27,000 per year on average) working exclusively on supervision of WaMu. *See* WaMu Report, *supra* note 7, at 17, Table 5. Examinations averaged about 150 days in length and were conducted by the equivalent of 20 full-time employees. *See id.* at __ (based on the data in Table 6).

[10] *See* Henderson & Tung, *supra* note 8, at 1040-45.

[11] *See* Steve Denning, *Big Banks and Derivatives: Why Another Financial Crisis is Inevitable*, FORBES, January 8, 2013 (available at http://www.forbes.com/sites/stevedenning/2013/01/08/five-years-after-the-financial-meltdown-the-water-is-still-full-of-big-sharks/)

[12] For an account of the S&L Crisis, see JAMES R. BARTH, THE GREAT SAVINGS AND LOAN DEBACLE (1991).

[13] For an estimate of the costs of the S&L Crisis, see Timothy Curry & Lynn Shibut, "The Cost of the Savings & Loan Crisis: Truth and Consequences," FDIC Banking Review (1995). For an estimate of the direct fiscal costs of the current financial crisis, see David Luttrell, et al., "Assessing the Costs and Consequences of the 2007-2009 Financial Crisis and Its Aftermath, Econ. Letter, Federal Reserve Bank of Dallas, vol 8 (2013), available at http://www.dallasfed.org/research/eclett/2013/el1307.cfm.

[14] Id.

phenomenon; historically, the US has experienced one about every twenty or thirty years.[15]

In response to the most recent bank failures, new systems of bank regulation have been proposed and, to some extent, implemented. New rules require banks to claw back from executives erroneously earned pay, based on a negligence standard.[16] Federal Reserve guidance effectively mandates restricted periods for performance-based compensation.[17] The so-called Volcker Rule, as currently proposed, would restrict the ability of banks to engage in proprietary trading, deemed too risky.[18] Several prominent academic commentators propose mandating inside debt compensation for managers.[19] Others propose regulators be empowered with the authority to prescreen bank activities to see whether they will improve social welfare.[20]

---

[15] There were large banking crises in (at least) 1819, 1837, 1857, 1873, 1884, 1890, 1893, 1907, 1933, the late 1980s, and the late 2000s. For a more general treatment of banking crises, see CARMEN M. REINHART & KENNETH ROGOFF, THIS TIME IS DIFFERENT: EIGHT CENTURIES OF FINANCIAL FOLLY (2010).

[16] *See* FDIC Rules Part 380.7 (12 C.F.R. Pt. 380) (available at http://www.fdic.gov/regulations/laws/rules/2000-9400.html#fdic2000part380.7)

[17] *See* Guidance on Sound Incentive Compensation Policies, 74 Fed. Reg. 55227, Docket No. OP- 1374 (proposed Oct. 27, 2009) (specifically stating the Federal Reserve's "expectation" that banks will utilize long-term compensation methods, including deferral); *see also* See Dodd-Frank Wall Street Reform and Consumer Protection Act of 20 10, Pub. L. No. 111-203, *available at* http://www.gpo.gov/fdsys/pkg/PLAW- lllpubl203/ content-detail.html (Congress requires that the appropriate Federal regulators shall jointly prescribe regulations or guidelines that prohibit any types of incentive-based payment arrangement, or any feature of any such arrangement, that the regulators determine encourages inappropriate risks by covered financial institutions - (1) by providing an executive officer, employee, director, or principal shareholder of the covered financial institution with excessive compensation, fees, or benefits; or (2) that could lead to material financial loss to the covered financial institution).

[18] *See* Title VI of the Dodd-Frank Act, Pub. L. 111–203, H.R. 4173 (2010). The Act defines "proprietary trading" as engaging as a "principal" regarding the trading account of a banking organization or other supervised nonbank financial entity in any transaction involving the purchase or sale of securities or commodities. *See* Act § 619 to be codified at 12 U.S.C. § 1851(h)(4).

[19] *See, e.g.*, Fred Tung, *Pay for Banker Performance: Structuring Executive Compensation for Risk Regulation*, 105 NW. U. L. REV. 1205 (2011) (proposing bank executive compensation contracts include securities tied to a bank's subordinated debt securities, since the price of these securities are tied to the downside risk of the bank, and therefore will give bank executives incentives to reduce risk taking to socially optimal levels).

[20] Eric A. Posner & Glen Wyle, *An FDA for Financial Innovation: Applying the Insurable Interest Doctrine to 21st Century Financial Markets*, 107 NW. U. L. REV. 1307 (2013) (arguing for an FDA-like administrative agency to pre-approval new financial instruments before they are offered to the public or used by banks or other nonbank financial institutions).

While apparently now discarded, there was a serious push for a tax on financial transactions[21] and a limitation on the size of banks.[22]

We are not optimistic that any of this regulation will prevent another banking crisis, and believe some of these reforms may actually *increase* risk in the banking sector. We make our argument as follows. First, we describe how current and proposed banking rules fit into familiar categories of regulation in other areas—that is, ex ante and ex post regulation—and consider the general costs and benefits of these approaches. We also identify a third type of regulation, the examiner's on-site evaluation of bank risk, which we call the "regulatory veto." While in other areas of economic activity we generally rely either on ex ante regulatory approval (e.g., drugs approved by the FDA or rules prescribed by the EPA) or ex post litigation (e.g., products liability law), in banking there is significant regulatory power that lies between these two poles. This is an underappreciated, but core, function of banking regulation.

Second, we consider ex ante and ex post regulation in banking, and find that they are either incapable of containing systemic risk and preventing future financial crises or unlikely to strike the optimal balance between regulation and wealth creation. The failure to contain risk arises from the gameability of ex ante regulation, as banks and bank shareholders can often assemble the same overall risk/reward profiles via alternative means. The rapid pace of financial innovation, and the inability of regulators to anticipate ahead of time all undesirable activities and activity levels, means that lawmakers are often plugging one of many holes in the dyke, to little positive effect.

To the extent that ex ante regulation does constrain behavior, we argue that it is unlikely to be particularly helpful. The reason is that regulatory prescreening fails to capture information about the social value and costs of bank activities. Ex ante regulations, if they work, do not allow parties to act fully upon their private information. For instance, a one-size-fits-all approach to bank size will necessarily be over-inclusive (limiting good banks from socially beneficial activities above the portfolio limits) and under-inclusive (allowing bad banks to engage in socially

---

[21] In 2009, various proposals were introduced in the House to take securities transactions. For instance, Rep. Peter DeFazio (D-Oregon) and Sen. Tom Harkin (D-Iowa) proposed a tax of 25 basis points on securities transactions. See Daivd Rogers, "Nancy Pelosi pushes global financial fee," POLITICO, Dec. 3, 2009, available at http://www.politico.com/news/stories/1209/30200.html.

[22] *See, e.g.*, Jonathan R. Macey & James P. Holdcroft, Jr., *Failure Is an Option: An Ersatz-Antitrust Approach to Financial Regulation*, 120 YALE L.J. 1368 (2011) (proposing to limiting any "financial institution from amassing liabilities in an amount greater than five percent" than the FDIC insurance fund for deposits).

undesirable activities up to the portfolio limits). This problem of lack of information (about both costs and benefits) plagues existing proposals to regulate bank activities directly, such as proposals for smaller banks, limits on amounts and types of executive compensation, and insurance requirements, in the form of capital cushions. Further, the regulator's information may improve over time, a benefit which is lost when both banks and the regulator are constrained by ex ante edicts.

Finally, ex ante regulation is plagued by public choice problems. The dynamics of congressional lobbying and the regulatory agency revolving door likely favor established institutions and tend to retard both new entry and economic growth. Politically powerful banks rather than competent ones may be able to win concessions from regulators. Other industries, such as securities underwriting and commodities trading, may seek to limit competition by prohibiting bank entry into those areas.

Ex post regulation is no panacea, either. While legal action is available to both private parties and the government to seek compensation for losses caused by harmful banking practices, these suits are unlikely to be effective deterrents. As a practical matter, such suits are extremely rare. In terms of government suits, this could be because of the political influence large banks have always had owing to the fact that the government relies on them to create wealth.[23] In terms of private suits, this may have to do with the complexity of the cases, the expected litigation costs of making a suit, and the difficulty of proving a breach of a duty in this setting.[24] But the larger failing of ex post regulation is judgment-proofness. Where potential losses are very, very large, as

---

[23] The decision to not prosecute Goldman Sachs for its role in the Financial Crisis was decried by some commentators as an example of cronyism between large Wall Street financial institutions and regulators, occasioned by the revolving door between regulated firms and government. *See, e.g.*, Peter Schweizer, *Why Goldman Sachs, Other Wall Street Titans Are Not Being Prosecuted*, DAILY BEAST, Aug. 14, 2012, *available at* http://www.thedailybeast.com/articles/2012/08/14/why-goldman-sachs-other-wall-street-titans-are-not-being-prosecuted.html.

[24] The barriers for civil suits by government and private parties are substantial. Banks generally do not owe a fiduciary duty to their clients, and contracts between the bank and its customers are filled with disclaimers that limit the bank's liability. *See, e.g.*, Peter J. Henning, *Is That It for Financial Crisis Cases?* NY TIMES DEALBOOK, Aug. 13, 2012, *available at* http://dealbook.nytimes.com/2012/08/13/is-that-it-for-financial-crisis-cases/?_php=true&_type=blogs&_r=0. Plaintiffs also have to show scienter, which can be extremely difficult in these cases. *See id.*; *see also* Ernst & Ernst v. Hochfelder, 425 US 185 (1976). This is especially the case when both parties are sophisticated entities, as was the case in most of the allegedly fraudulent transactions leading up to the Financial Crisis. *See id.*

they are for a banking crisis, most defendants are likely to be bankrupt and unable to pay any liability assessed. Mandatory insurance is a typical way to mitigate the judgment-proofness problem, often used in other industries, but this may be ineffective in the case of bank failures because of the systemic nature of the financial system: the entire industry may fail together, swamping the resources of any mandatory insurance system.

In any event, so long as banks face expected losses from wrongful activities that are lower than the externalized costs, banks will engage in too much banking activity and take on too much risk. In short, there is some value to regulatory delay because the quantity and quality of information held by regulators, and thus the quality of regulatory decisions, is increasing in time, but waiting too long is dangerous because the risk of judgment-proofness is also increasing in time.

Third, we point out an unusual feature of banking regulation, which we believe is designed to ameliorate the problems caused of both ex ante and ex post regulation by taking an intermediate position. This is the institution of the bank examiner, who is authorized to veto or shut down a bank's operations based on an assessment of the bank's type and amount of activities on an ongoing basis. We call this feature a "regulatory veto," since it is a determination about whether existing activities should and can continue, made after they have commenced but before they have completed. The regulatory veto is intended to strike a balance between ex ante and ex post regulation, trying to avoid the informational problems of the former and the it's-too-late problem of the latter.[25] For instance, it appears that examiners at Washington Mutual successfully identified excessive risk concentrations in some asset classes;[26] forcing an unwinding of some these positions could, in theory, have at least limited the magnitude of Wamu's failure and its effect on the wider banking system.

Finally, despite the advantages of this intermediate approach, we describe a fundamental flaw of the regulatory veto, and point to this as a major source of the inefficient regulation of

---

[25] For instance, if a bank decides to engage in lending concentrated in a particular industry or geography, the regulator could decide that the social risks are too high, and therefore demand the bank shut down the particular lending or the entire bank. The veto is exercised after a decision about whether to engage in the activity has been made by bankers and after regulators can observe social cost of those activities. A more efficient allocation of resources is possible because shareholders can act on their signal of prospective firm value, and regulators can act based on an ex post signal of firm value and riskiness.

banks. While existing regulatory reforms focus on bad decisions or incentives of bankers, we argue that it is actually a game-theoretic problem that keeps the regulatory veto from functioning correctly: the veto game between the bank and the regulator is an ultimatum game in which the bank is the first mover. As we show below, since the bank gets to decide upon the activity before the examiner can veto it, the bank effectively presents the regulator with a take-it-or-leave-it offer. So long as overall social welfare is not sufficiently negative, the regulator will weakly prefer not to shut down an overly risky activity, since firm production yields a net societal benefit of zero and the costs of shutdown are positive and significant. The result is socially-inefficient, excess amounts of bank activity – not as high as without the regulatory veto, but still potentially far above the social optimum.

Thus, the regulatory veto model is a vital yet flawed component of banking regulation. Therefore, our prescription for bank regulation reform is to try to improve the dynamics of the bank-regulator game, so that making take-it-or-leave-it offers will be more costly for banks. Social welfare outcomes (though not firm profits) could be improved by giving the regulator an arsenal of credible threats to veto bank activity; threats may be made credible by altering regulator incentives[27] or enlarging the regulator's permissible action set. Ex ante regulations that limit the costs of bank shutdown (such as "living wills") make the regulatory veto more efficient. We outline those areas we think most ripe for improvement.

II. THE BANKING SYSTEM: A PRIMER

This section lays the groundwork for the normative analysis in Section III. We first discuss the generally accepted economic model of bank wealth creation, in which banks turn short-term liabilities (such as bank deposits) into long term assets (such as mortgage loans). The mismatch between the maturities of liabilities and assets carries with it particular vulnerabilities that, given the importance of banking to the economy, may make some regulation desirable. We also describe how this model, and the rationale for regulation, extends to non-banks financial institutions that perform some traditional banking functions.

---

[27] *See, e.g.*, M. Todd Henderson & Frederick Tung, *Pay for Regulator Performance*, 85 S. CAL. L. REV. 1003 (2012) (proposing to compensate bank regulators with securities mimicking the stock price and debt price of the banks they supervise in order to give the regulators upside and downside incentives from bank performance).

Next, we describe the current system of banking regulation. In our taxonomy, banking regulation is divisible into ex post and ex ante measures. Ex ante measures or those in which regulators require or prohibit certain bank activities (or levels of activity) before banks have undertaken any activity. Ex post describes the case where the regulator or legal system waits until after the banking activity has occurred and any harm has been realized before seeking remedies (or imposing punishments) against the bank and its managers. One form of bank regulation, that of the bank examiner and the regulatory veto, finds a middle ground between ex ante and ex post regulation.

Finally, we describe banking reform proposals that have surfaced in recent years. These proposals include bank size limits, a financial transactions tax, a financial FDA, and greater penalties on bank managers. We classify these reforms into our ex ante/ex post taxonomy. The ex ante/ex post classification aids our economic analysis of banking regulation, which we undertake in Part III.

A. Banking Basics – Long-short mismatch, runs, and maintaining confidence

The effective and efficient regulation of banks is crucial to a well-functioning economy. Banks perform the crucial role of intermediating short and long term financial investments – in effect creating the credit necessary for both businesses and households to function. This role, in addition to other ancillary functions of banks (such as operating payment systems and trading in securities), makes banks systemically important in that the failure of many (or even one very large) bank is generally thought to have significant economic consequences beyond just the bank's security holders and counterparties.

A problem with banks, though, is that the very feature that allows them to create value also makes them vulnerable to insolvency due a decrease in asset values – or even for no reason at all. Following the famous Diamond & Dybvig model of banking, the central business model of banks (and so-called shadow banks, who are financial firms that undertake the same sort of strategy) is that banks borrow on short term maturities at a low interest rate, lend at long term maturities at a high interest rate, and pocket the difference in interest rates on the two.[28] Figure 1, below, depicts a bank that takes various forms of primarily short term liabilities

[28] *See* Douglas W. Diamond & Phillip H. Dybvig, *Bank runs, deposit insurance, and liquidity*, 91 J. POL. ECON. 401 (1983) (providing a model of banking stability given asset and liability mismatch, and of bank runs and financial crisis in which even solvent banks are fragile under the stability model).

(such as deposits, repurchase agreements) and converts them primarily into long term assets (such as mortgages and term loans to businesses); revenues from the assets are used to pay interest and principal on the short term liabilities.

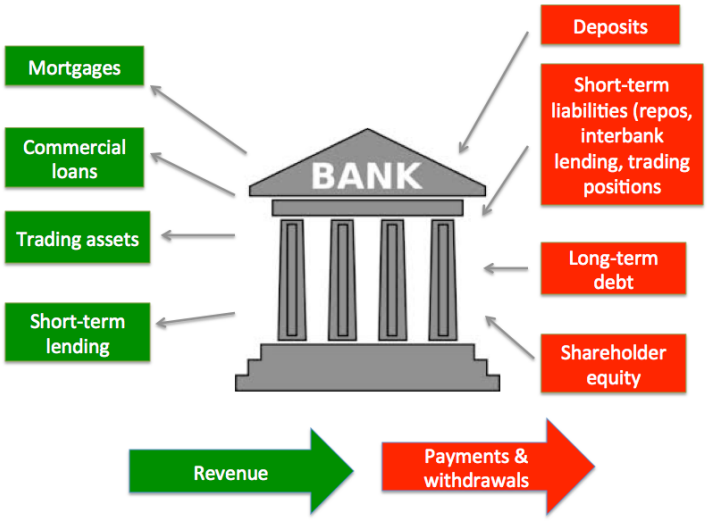

Figure 1: A Bank

Consider a simple bank that makes loans funded by the accounts of depositors. The prototypical bank deposit is a "demand" deposit, meaning that the depositor has the right to withdraw her money whenever she desires; this is the shortest possible maturity of borrowing, in that it is due and payable at any time upon the demand of the depositor. Depositors like such flexibility, as they may suddenly need their money to cover unanticipated expenses, such as medical bills or other unforeseen liabilities. Nevertheless, despite the promise to repay the depositor on demand, the bank is able to turn this deposit into long term money that it can lend to other individuals or businesses. So if the depositor deposits $100 in the bank, the bank may turn around and immediately lend out this $100 to a business or homebuyer for a much longer term (the typical mortgage, for instance, has a 30-year maturity). In this way, the liquidity needs of the depositor and the long term capital needs of the borrower can be met; the bank has effectively turned the depositor's $100 of short term cash into $100 of long-term assets. In the language of macroeconomics, the bank has increased the supply of money by $100, from $100 to $200, allowing additional investment and growth in the economy.[29]

[29] *See* N. GREGORY MANKIW, PRINCIPLES OF ECONOMICS 6th Edition at 630 (2012)

What happens, however, when the depositor wants her money back? In the above example, the bank would be unable to call in its loan to the borrower since it is not yet due, and the bank would therefore be unable to pay its liabilities as they come due – the bank is insolvent, in other words. There are two things that banks do, in tandem, to avoid such an outcome. Banks, first of all, diversify across depositors, and, second, they do not lend out all of the money that depositors deposit. This amount that banks hold, rather than lend, is known as "reserves." The more diversified the depositor base, the lower the reserves that a bank must hold, and the more credit it can extend to borrowers. This leads to the system of "fractional reserve banking," as it is known, in which banks must hold only a fraction of their deposits as reserves.[30] Notably, this prudential measure arises even without the existence of government regulation. No bank that lent out too much of its deposits would stay in business long. But, notwithstanding banks' incentives to remain solvent, national and international laws regulate the amount of cash banks must hold on hand and limit the amount of leverage banks can take, as described below.

Apart from diversifying depositor accounts, banks maximize the amount that they can safely lend by using experience and sophisticated models to predict withdrawals, thereby saving a bare minimum to meet demand deposit requests. Additionally, if banks find themselves with too few reserves, banks can typically borrow reserves from another entity (such as other banks, financial firms, and even the government), sell assets, raise more deposits, or issue securities. A problem with these stopgap measures is that they may not work in a systemic crisis: other institutions may be unwilling to lend; asset prices may have already declined, with further sales increasing the downward price pressure; depositors (or other lenders) may fear losing their investments; and shareholders would be unwilling to commit more capital if the bank's assets are less than its non-equity liabilities.

But low reserves make banks more susceptible to a bank run. A run happens when depositors lose confidence their money is available or will be available in the future, causing them to demand their money back immediately. As viewers of "It's a Wonderful Life" or "Mary Poppins" know, a bank run could wipe out a *solvent* bank that does not have the instantaneous ability to convert its illiquid assets to cash. In a sort of self-fulfilling prophecy, if one depositor believes that other depositors are likely to demand their deposits back, then it makes sense for her to demand as well, since otherwise, if the reserves are inadequate,

30

she will, at best, not immediately have access to her account, and, at worst, lose some or all of her investment. Thus, in what has been termed a rational panic,[31] all depositors may rationally demand their deposits back at the same time, even though no fundamental change has occurred at the bank.

While rational panics and bank runs may occur for no substantive reason at all, deficiencies in a bank's capitalization make bank panics and bank failure more likely. Certainly, the amount of reserves on hand is one such important variable: as the likelihood of being unable to meet withdrawal needs increases, depositors and other creditors will be less willing to keep their money with the bank. Similarly, as the value of a bank's assets (its mortgages and loans) declines, due to, say, a deterioration in the real estate market, the likelihood of substantive insolvency increases as well. Conversely, as new equity is injected into the bank, creating a cushion for depositors and other creditors, the likelihood of substantive insolvency lessens, and bank runs become less likely.

The solution, developed in response to the banking crisis that triggered the Great Depression, is government insurance. To reduce the fear of a bank run, and thus to decrease the cash necessary to be held on hand and therefore to increase the speed at which wealth can be created, the government stands behind bank deposits, promising to pay if the bank does not.[32] The primary mechanism is guarantees of deposits up to $250,000 by the Federal Deposit Insurance Corporation (FDIC), which charges member banks a risk-adjusted premium.[33] The FDIC acts on the assumption that if depositors know that they have assurances from the government, they have no need to pull out their money.

Yet the solution to bank runs creates another problem. Let the depositors know their deposits are secure, and they will care little about the prudence of their bank's lending operations. The government guarantee satisfies their worries about nonpayment, no matter how badly the bank performs.[34] This is the familiar

---

31

32 These concerns are not present, or are present to a much less degree, in a world in which banks issue bank-specific notes of general circulation. For a discussion of a banking system comprised of "private money," see, for example, Bruce D. Smith & Warren E. Weber, *Private Money Creation and the Suffolk Banking System*, 31 J. MONEY, CREDIT AND BANKING 624 (1991).

33 *See* Federal Deposit Insurance Corporation website, The Deposit Insurance Fund, *available at* http://www.fdic.gov/deposit/insurance/index.html.

34 Current rules limit the amount of deposit insurance to $250,000, but individuals can evade this limit by holding accounts of up to this amount in multiple banks, each of which is insured for the maximum amount. *See* FDIC FAQs, http://www.fdic.gov/deposit%2Fdeposits/ ("FDIC insurance covers all deposit accounts, including checking and savings accounts, money market

moral hazard problem, and it inevitably leads to a less than optimal amount of private monitoring of bank risk taking. It therefore falls on the government, as it falls on every guarantor, to take steps to monitor the bank so that it does not engage in risky activities that could trigger claims against it under the guarantee. Out of the government's need to control moral hazard by its banks, the system of bank regulation has been born.

Banks are also dependent on non-deposit funding. Banks may borrow from other banks and financial institutions, establish trading positions that may incur liabilities in the near future, issue longer term certificates of deposits, sell bonds, and undertake new stock offerings. These stakeholders do not benefit from FDIC insurance – in fact, because the government gives priority in bankruptcy to FDIC-insured accounts, one could say that they lose from it – and hence their incentives to pull credit from a troubled bank remain strong. In a slower-moving analog of a bank panic, these sorts of various bank stakeholders may refuse to invest new funds and, further, refuse to roll over old debts as they come due. Lehman Brothers, for instance, found itself unable to roll over its short term liabilities, largely securities repurchase agreements ("repos"), even though Lehman remained balance-sheet solvent.[35]

Thus, the problem of bank runs poses a problem not just for actual banks, but for the variety of financial firms that undertake bank-like activities, the "shadow banks." Shadow banks are dependent upon short term liabilities to fund long term assets, and therefore face the very same vulnerabilities as do actual banks. Money market funds, some hedge funds, structured investment vehicles, and some investment banks (such as Lehman) borrow in short term credit markets to purchase longer term assets. Some of these shadow banks, such as Lehman and Reserve Primary Fund (the oldest money market mutual fund) did experience failures that were as potentially damaging to the financial system as were the actual bank failures that occurred.

Given the fragility of the banking and shadow banking systems, as well as the important part these institutions play in the economy, it is unsurprising that they are highly regulated, and becoming more so. There are a variety of regulatory tools and approaches to optimizing the amount and type of bank regulation,

---

deposit accounts and certificates of deposit. The standard insurance amount is $250,000 per depositor, per insured bank, for each account ownership category.").

[35] According to Lehman Brothers's last 10-Q, filed for the second quarter of 2008, Lehman's stockholders had positive equity of $26 billion. *See* Lehman Brothers Holdings Inc. 2Q 2008 10-Q (available at http://www.sec.gov/Archives/edgar/data/806085/000110465908045115/a08-18147_110q.htm).

including addressing the problem of bank runs. We consider these in the next section.

B. *B*anking Regulation Basics

There are two central issues in banking regulation. The first is the moral hazard problem created by FDIC insurance, itself an attempt to solve the bank run problem. Depositors need not worry about whether the bank invests prudently, and whether the bank is adequately capitalized. Instead, the FDIC and the federal government bear that risk. In other words, the solution to alleviating one sort of risk – the rational panic – actually induces more risk-taking behavior on the part of banks because more of the banks' costs can be externalized onto third parties. Short of a radical change to our entire banking system,[36] this government guarantee generates a social liability that means even many small-government types wholeheartedly support a robust and vigorous system of banking regulation.[37]

The second is that banks are thought to generate significant externalities, both positive and negative, such that a bank's owners and managers will have incentives that are not always in line with the best interests of society. On the positive externality side, banks play a central role in the creation of credit and the expansion of the money supply. The government is, to an extent, dependent on banks to help the economy run and expand; and if banks, as they were in the recent Financial Crisis, are unwilling to lend, the government's plans to kick-start the economy will be greatly hampered.[38] Hence, a significant problem can be, at times, banks taking on too little risk and engaging in too little banking activity.

---

[36] If bank runs are not a systemic problem, then one could imagine a return to the private-banking model that prevailed for much of US history. Individual banks would issue currency whose value would depend on the market's evaluation of the riskiness. *See* Thomas L. Hogan, "Competition in Currency: The Potential for Private Money," Cato Policy Analysis 698, available at http://www.cato.org/publications/policy-analysis/competition-currency-potential-private-money; see also DAVID GLASNER, FREE BANKING AND MONETARY REFORM (2005). We view the likelihood of such a radical change, whatever its merits, as vanishingly small.

[37] *See, e.g.*, Richard A. Epstein & M. Todd Henderson, *Do Accounting Rules Matter? The Dangerous Allure of Mark to Market*, 36 J. CORP. L. 513 (2011) ("In this environment, the rationale for banking regulation, which even small-government types support, runs as follows.").

[38] *See, e.g.*, Stephen Gandel, Study: *Bank bailout didn't boost small business lending*, CNNMONEY.COM, November 14, 2012, *available at* http://finance.fortune.cnn.com/2012/11/14/bank-bailout-lending/.

On the negative externality side, bank failures have consequences that may extend beyond the bank's shareholders and depositors. Following a relatively conventional account of systemic risk in the banking system, bank failure causes external harms by reducing confidence in payment systems, causing losses to counterparties who cannot properly calculate counterparty credit risk, and depressing asset prices as the bank scurries to offload its assets, which can lead other banks into crisis (sometimes referred to as "financial contagion").[39] More prosaically, losses to bondholders and other creditors are externalities from the bank's shareholders' point of view.

Thus, any system of bank regulation must attempt to encourage a robust amount of banking activity – credit being the lifeblood of a modern economy – while not allowing banks and bank shareholders to run totally rampant, maximizing profits at the expense of society as a whole.

Before considering some of the regulatory techniques designed to solve the problems created by deposit insurance and other banking externalities, it is important to note that all of them can be thought of as regulating either the amount of banking activity ("activity level regulation") or about how careful the bank is in its activities ("due care regulation"). These are the familiar objects of tort law, and therefore the literature on optimal regulation of accidents can be a useful analog in evaluating the efficacy of banking regulation.

At all levels and among all regulators,[40] banking regulation is designed to ensure the "safety and soundness" of banks.[41] Safety and soundness are related, but different concepts. Safety is about ensuring that depositors have a safe place to put their money where they know it will not be squandered or lost due to fraud, errors of judgment, or bad luck. Safety, or the appearance of

---

[39] For an example of the application of this argument, see VIRAL V. ACHARYA, MATTHEW RICHARDSON, STIJN VAN NIEUWERBURGH, AND LAWRENCE J. WHITE, GUARANTEED TO FAIL: FANNIE MAE, FREDDIE MAC, AND THE DEBACLE OF MORTGAGE FINANCE (2011).

[40] Although not relevant to our analysis, it should be noted that for a variety of reasons, the United States has a dual banking system, meaning there is bank regulation at both the state and federal level, and banks generally have a choice between them. In addition, banking activity can take place through either a "bank" or a "thrift," each of which has a different regulator at the federal and state level. This means there are at least four choices, which generates significant amounts of regulatory competition. This choice has traditionally made a difference on a number of dimensions, such as: reserve requirements, capitalization standards, and branching rules. For a discussion of the history and structure of US regulation of banks, see RICHARD S. CARNELL ET AL., THE LAW OF BANKING AND FINANCIAL INSTITUTIONS 2–34, 62–66 (4th ed.) (2009).

[41] *See id.* at 251-267.

safety, reduces the threat bank runs, which benefits everyone. Soundness is about ensuring banks are well managed and maintain adequate levels of capital against losses. Although these have been the pillars of banking regulation since the Founding, there are many ways in which regulators can try to achieve the optimal levels of safety and soundness, and these have changed over time.

1. Prudential Regulation

From the beginning of the Republic and for hundreds of years, safety and soundness regulation focused primarily on substantive regulation of banking practices, what is often called "prudential regulation." Federal and state laws limited entry into the banking business, defined the types of businesses and activities banks could engage in,[42] required approval for "fundamental changes" to bank ownership or activities,[43] and limited lending activities, including capping loans to any one borrower,[44] specifying levels and types of interbank lending,[45] and dramatically restricting lending to bank executives or other insiders.[46]

The conceit of this regulatory model is to impose a one-size-fits-all rule restricting banks from doing specific things that might cause their depositors losses or lead to reduced stability in the banking system (that is, cause losses for depositors at other banks). For purposes of perspective, it is interesting to note that this mode of regulation – which we term a form of "ex ante regulation" – used to also be the norm for corporation law generally. Historically, non-bank businesses had to seek governmental approval before they could engage in any type of business, and corporate charters were bespoke, in that they restricted firm activity to enumerated areas. This system of ex ante regulation of corporate formation died over a hundred years ago as New Jersey and then Delaware liberalized chartering requirements so that businesses could engage in any business activity, subject only to ex post regulation, such as litigation.[47] Rules forbidding deals with insiders also faded in favor of a disclosure and approval

---

[42] *See id.* at 107-128.
[43] *See id.* at 86-91; *see also* 12 U.S.C. §§ 214a, 215, and 215a.
[44] *See* CARNELL ET AL., *supra* note 40, at 296-300.
[45] *See id.* at 302-304.
[46] *See id.* at 304-307.
[47] *See, e.g.*, Charles M. Yablon, *The Historical Race Competition for Corporate Charters and the Rise and Decline of New Jersey: 1880-1910*, J. CORP. L. 324 (2007).

regime.[48] The only remaining form of pervasive ex ante regulation of non-bank activities is in certain business activities, such as pharmaceuticals or energy production, where there might be potential harms that are difficult to remedy ex post. Even potentially dangerous consumer products are not required to get a governmental approval before they are sold; we rely on the ex post tort system of products liability to optimize care and activity levels.

In banking, however, ex ante approval for activities persists. For example, section 24(7) of the National Bank Act provides that banks—called "associations"—are permitted to engage only in the "business of banking" and any businesses "necessary" to carry out the business of banking.[49] The power to construe these ambiguous terms is delegated to the various regulatory agencies that oversee banks, such as the Office of the Comptroller of the Currency (OCC). The regulators have generally taken a permissive attitude, allowing banks to enter into many business activities tenuously related to banking. For instance, the OCC has opined, over the objections of incumbents in the particular industry, that banks can offer travel agency services, sell annuities, provide Internet transaction services, and many other things hardly "necessary" to the business of banking.[50]

As in the regulation of corporate purpose, such ex ante regulation is fraught with problems. Regulators deciding whether to allow a particular bank to offer travel agency services or sell securities, for instance, do not know how efficient or effective the bank will be at offering the service, how much customer demand there will be, whether customers will be happy with the services provided, and what the potential risks are from the activity. To be sure, regulators have some experience with banks in general and, perhaps, the activity sought to be done by the bank, but they have less information about private and social costs and benefits than if they waited until the bank started offering the service and then re-evaluated. The quantity and quality of information is increasing in time.

---

[48] *See id.* The modern Delaware rule permits transactions with insiders under certain circumstances, such as disclosure of the conflict and a ratification by disinterested shareholders. *See* D.G.C.L. § 144.

[49] 12 USC 24(7). 12 U.S.C. § 92, permits national banks to engage in insurance agency activities if they are are "located and doing business in any place the population of which does not exceed five thousand . . . ." The purpose of this provision, added to the National Bank Act in 1916, appears straightforward: to provide people living in small towns with access to insurance agency services which might otherwise be unavailable or available only on non-competitive terms.

[50] See CARNELL ET AL., *supra* note 40, at 127-173.

There are other problems. Even if regulators had good information about the costs and benefits of a particular type of activity on average, it is possible, perhaps even likely, that the net of these differs significantly across different banks, making a one-size-fits all approach both over and under-inclusive. While Bank 1 might be able to engage in a particular activity in a way that generates significant social welfare, Bank 2 may not. Crucially, the regulators may not know this before Bank 1 and 2 are able to demonstrate it through actual practice.

Another problem is that the decision on whether to permit a particular type of activity will be influenced by factors unrelated to the efficiency or safety and soundness questions. Regulatory capture is a well-known and significant phenomenon, meaning regulators may be too willing to approve activities in some cases, while too willing to deny them when powerful competitors use their influence. Importantly, these problems, which plague regulation at any time, may be more significant for ex ante regulation, since conjectures are easier to counter or disprove than facts. One need only observe the large number of bank frauds in and general mismanagement of thrifts during the Savings & Loan Crisis of the 1980s, as well as the shady banking practices in the mortgage industry more recently, to see the perils of relying too heavily on an ex ante screening mechanism to keep banking standards high.[51]

Finally, often the real risk for banks and from banking arises not from forays into unrelated fields, like insurance or travel services, but rather from core banking activities. While a small local bank might be put in peril by over-lending to a particular individual, chicanery by insiders, or unsuccessfully entering into unrelated businesses, in the world of large, modern banks, systemic risk is much more likely to arise from bad investment decisions at the macro level in the banking industry as a whole. As a case in point, Washington Mutual's losses stemmed not from exotic derivatives or speculative side-bets, but rather from its core home lending business.[52]

---

[51] *See supra* note __.

[52] *See* Washington Mutual 2007 10-K filed February 29, 2008 at 16-17 (Wamu's deterioration was "primarily the result of significant credit deterioration in the Company's single-family residential mortgage loan portfolio and significant disruptions in the capital markets, including a sudden and severe contraction in secondary mortgage market liquidity for nonconforming residential loan products. These conditions also contributed to the impairment of all goodwill associated with the Company's Home Loans business near the end of 2007.") (available at http://www.sec.gov/Archives/edgar/data/933136/000104746908002083/a2182890z10-k.htm)

So although these ex ante rules and requirements are still on the books, they are much less an important part of systemic risk regulation today. Another form of ex ante regulation, risk-based capital requirements, plays an increasingly important regulatory role. The logic of moving from regulation of business activities to capital requirements is straightforward. The significant information problems faced when determining the risk from specific activities ex ante go away when banks are free to do whatever they want, so long as they maintain sufficient "insurance," in the form of minimum levels of equity capital that can be brought down without endangering the health of the bank's creditors.

2. Capital and Reserve Regulation

Absent banking externalities, there would be no need for government-mandated capital regulation, since banks that did not keep enough cash on hand would be driven out of business, and banks that do not maintain an adequate amount of equity to cushion the losses of depositors and other creditors would not be able to raise funds. But the externalization problem – bank runs, government insurance, and other banking externalities – means the laissez faire equilibrium bank behavior would not be at the social optimum. As noted above, private monitoring is inefficiently low given government insurance, which means that the market, if left alone, will reach an equilibrium that is privately optimal (to shareholders and bank managers), but certainly not optimal from a societal perspective. It is important to note that the optimal reserve and capital amounts will vary across firms, depending on their activities, experiences, skills, and other variables.

To solve the problem of inefficient market pressure, government estimates the optimal point and then requires all banks maintain a minimum amount of regulatory capital. Because the optimal reserve capital cannot sensibly be one number for large and small banks, regulators determine the amount of regulatory capital as a ratio of capital to firm size, usually determined by assets. For many years, the typical ratio was known as the "leverage limit," which was the total bank capital to total bank assets, where, roughly speaking, bank capital is a measure of shareholder equity, while total bank assets is the bank's entire portfolio of loans, mortgages, cash on hand, and any other investments.[53] To satisfy the regulatory requirement, banks had to

---

[53] *See* 12 C.F.R. §§ 3.6, 6.4(b)(2)(iii); *see also* CARNELL ET AL., *supra* note 40, at 256-257.

keep at least four percent of capital to assets.[54] Although the ratio concept is straightforward, it is complicated to determine what precisely goes into total capital and what liabilities on the balance sheet (and off) should be considered in determining the numerator and denominator.

Putting these complexities aside, however, one can easily see the over- and under-inclusiveness problem with using a single, ex ante leverage ratio to determine the optimal bank capital. Four percent will be too low for some banks, while it will be too high for others. To be sure, banks with greater risks may voluntarily hold more than four percent, but because of government insurance, whatever amount they hold will be less than they would hold if subjected to efficient market forces. So, although ex ante capital regulation has the benefit of not micromanaging bank decision making and substituting the judgment of bureaucrats for bankers, it suffers from a significant problem that faces all ex ante regulation.

The one-size-fits-all problem became a significant problem during the 1980s, as banks became more heterogeneous in the type of risks they were taking on.[55] Greater sophistication and diversification of bank activities necessitated an approach to regulatory capital that recognized this change. To address risk heterogeneity, in 1988 the Bank for International Settlements in Basel, Switzerland proposed a regulatory requirement of eight percent capital to risk-adjusted assets.[56] This was a nod in the direction of a capital policy tailored to individual bank risks and quality, but because it was executed ex ante it necessarily maintained a one-sized-fits-all approach – eight percent capital for certain risks might be too much, too little, or just right, depending on the bank. Most countries adopted the Basel Accord (known as "Basel I"), including the United States.[57]

Other than the inclusivity problem noted above, the eight-percent rule seemed to work well for many years, allowing some tailoring of capital requirements to bank risk. But by the early 2000s, regulators concluded the rule was ill suited for large banks operating across borders and in highly sophisticated markets. Accordingly the BIS promulgated a new capital reserve regime in 2004. Basel II made many changes,[58] but the core was permitting

---

[54] *See id.*

[55]

[56]

[57] *See* Basel Capital Accord (Basel I), Basel Committee on Banking Supervision, July 1988, *available at* http://www.bis.org/publ/bcbs04a.htm

[58] The three pillars of the revised Basel accord were capital (sole focus of Basel I); market discipline (mandatory disclosure of bank's condition); regulatory supervision. *See id.*

large, sophisticated banks to tailor their risk-based capital reserves by either: (1) an enumerated asset-specific risk assessment;[59] or (2) the bank's internal risk assessment models.[60] For instance, for the first approach, sovereign debt did not count toward liabilities, while 20 percent of exposure to OECD banks counted, 50 percent of residential mortgages, and 100 percent of consumer loans and other unsecured debt.[61] Again, this adjustment to the ex ante regulatory regime was an attempt to tailor risk, but clearly shows the problems of blanket rules, as well as attempts to predict the risk of particular bank activities as general categories. For one, greater weighting may encourage regulatory arbitrage, as in the use of credit instruments to move certain risks off of bank balance sheets. In addition, as the housing bubble and ensuing financial crisis shows, exposure to residential mortgages was riskier than believed, and not all banks holding it were good ones or worthy of the same risk weighting.

The numerous bank failures of the past few years point to the problem of relying too heavily on ex ante capital requirements, especially ones linked to rating-agency-based assessments of risk or internal bank models. Capital, however defined, seems to be evidence of bank weakness after the fact, rather than a predictor of it in advance. If a bank makes bad or overly risky investments, it is not until those asset prices actually fall that the bank shareholder equity shrinks, making it then undercapitalized.[62] Moreover, ex ante regulatory approaches such as this may generate false confidence that disarms whatever market pressure would otherwise exist.

3. The Regulatory Veto

Perhaps out of recognition of these problems, the existing regulatory regime for banks includes an uncommon feature: regulators, called "bank examiners," have the power to shut down bank activities midstream if the regulators believe the social costs are too high. We call this feature a "regulatory veto," since it is a determination about whether existing activities should and can

---

[59] 100% of government; etc. tied to rating agencies. Whoops.

[60] Basel II also added in the concept of "operational risk," which simply meant adding in some additional capital as a cushion in the event of a failure in the bank's operations, computer systems, or the like.

[61] *See id.*; *see also* CARNELL ET AL., *supra* note 40, at 257-265 (showing how to perform capital calculations under the rules).

[62] This is assuming mark-to-market accounting. In book value accounting, where banks do not need to write down asset values, the bank could still remain adequately capitalized for regulatory purposes, even though in real terms it is not. *See* Epstein & Henderson, *supra* note 37, at __.

continue, made after they have commenced but before they have completed. After loans or investments have been made, bank regulators are authorized to veto a bank's operations based on an assessment of the social welfare of the bank's choice about its type and amount of activities. The veto can take the form of a modification of the way the business is being conducted or a wind-down and termination of the business.

For instance, if a bank decides to engage in lending concentrated in a particular industry or geography, the regulator could decide that the social risks are too high, and therefore demand the bank to stop lending, alter its origination standards, increase oversight of the portfolio, take on more capital against losses, or shut down the particular lending or the entire bank. The key virtue of the regulatory veto is that it is made after a decision about whether to engage in the activity has been made by bankers and after regulators can observe social cost of those activities. The idea is to find the point where the information about costs and benefits of the activity in question is at a high point relative to the expected cost of delay. A more efficient allocation of resources may be possible because shareholders can act on their signal of prospective firm value, and regulators can act based on an ex post signal of firm value and riskiness.

The topology of regulatory choices is shown on Figure 1.

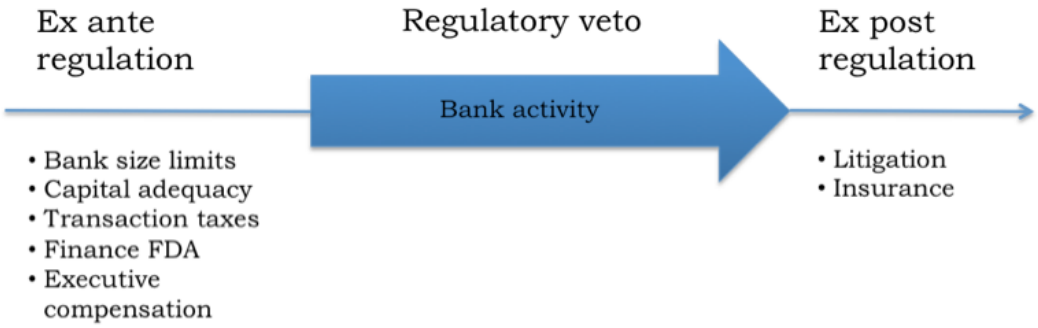

Figure 2: Topology of Regulatory Mechanisms

The regulatory veto is designed to allow timely and tailored regulatory intervention. It allows information about regulatory costs and benefits to be learned by regulators, who are watching banks closely. The presence of an onsite examiner is an opportunity for bank regulation to be different than other forms of regulation, which rely entirely on ex ante regulation, ex post regulation, or a mix thereof. Consider the regulation of driving. We try to make roads safe by using a mix of speed limits (ex ante rules) and tort suits (ex post litigation) to produce the optimal

amount of driving and care about driving. The blanket ex ante rules will, of course, be over- and under-inclusive. It would be a social welfare improvement if we could design and enforce rules based on the expected costs and benefits of each driver. A racecar driver could easily drive faster than the speed limit with less risk than a teenager driving less than it. However, it is simply too costly to write driver-specific rules, as the monitoring and enforcement costs would simply be too high. Narrowly tailored rules may be more efficient without considering the costs of enforcement, but when they are considered, rules of general applicability, despite their clunkiness, may be superior.

But if it is possible with reasonable cost to have rules designed to optimize the social costs and benefits of a particular actor, then regulation can be improved. In the speed-limit hypothetical, if the highway patrol could have an officer riding in every car, then it might be possible to create local regulation that would be more efficient. As discussed below, the examination process for banks is just that. Examiners sit at banks, examining each bank's activities for risk. If this process worked well, and we show it does not at present, it could allow regulators to maximize the value of regulation and therefore improve social welfare by reducing bank activities.

Some of the details of the examination process have been recounted in other work,[63] but it is worth summarizing some of the specifics here. In practice, bank examiners work full time monitoring large bank activities for compliance with regulations and established risk tolerances.[64] An examination occurs once per year for routine cases, and more often as needed based on regulators' views of the riskiness of the bank in question.[65] Many

---

[63] *See, e.g.*, Henderson and Tung, *supra* note 8.

[64] Several federal agencies supervise banks: the Office of the Comptroller of the Currency (OCC) supervises national banks; the Federal Reserve (Fed) supervises state member banks and bank holding companies, and the Federal Deposit Insurance Company (FDIC) supervises state non-member banks and FDIC-insured savings banks. *See* FDIC Report, *supra* note 8, at 463. The FDIC also has back-up supervisory responsibility for monitoring the condition of national banks and state member banks. In fulfilling these responsibilities, it works with the other two federal regulatory agencies. Under the Financial Institutions Reform, Recovery, and Enforcement Act of 1989 (FIRREA), it also has back-up authority to examine thrift institutions as well. State banking departments supervise state-chartered banks. *Id.*

[65] The frequency of examination varies by agency and over time. For instance, the National Bank Act of 1864 mandated that the OCC examine all national banks twice a year but allowed an extension to three examinations every two years. This policy stood until 1974, when the OCC moved toward off-site examinations using statistical methods, and the average examination schedule was more like 18 months. With the passage of the FDICIA, on-site examinations were required by law. By the late 1980s, resident examiners were placed in the

thousands of person hours are typically spent conducting the examination of large banks.[66] Examiners judge the credit quality of each asset, that is, loan or investment that has been made. The assessment is made based on discussions with loan officers and bank managers. Examiners also review loan portfolios as a whole for issues such as concentration risk, violations of legal rules, and deviations from bank loan and underwriting policies. They also judge other aspects of bank activity, such as the affairs of subsidiaries and affiliates, litigation risks, off-balance-sheet activities, and activities of insiders. Importantly, they do this all after lending decisions have been made and after loans have experienced some real-world conditions to put their risk into context.

After the on-site examination is completed, the examiner first presents a report on bank activities to management. Bank officials are able to comment on the report and offer to fix any deficiencies. The report is also taken to the bank's board of directors, which can promise to take corrective action to preempt regulatory action to shut down bank activities that threaten systemic risk.[67] For example, the decision to drop a bank's "CAMELS" rating from 2 to 3 (moving the bank from "fundamentally sound" to indicating "some degree of supervisory concern") precipitates formal or informal actions, which include obtaining the bank's written commitment to take corrective action.[68]

For bank activities that are determined to pose excessive risk or are otherwise not compliant with banking rules and regulations, examiners have enormous power to change bank practices. The regulatory pressure can be either explicit or implicit, backed by a

---

largest multinational banks, and by the 1990s, larger regional banks also got resident examiners. *See* FDIC Report, *supra* note 8, at __. Similar changes were also true of FDIC and Treasury examinations. FDIC examination periods varied from one to three years, depending on the CAMELS rating of the bank in question. Like for the OCC, however, the FDICIA mandated "annual on-site examinations of all banks except highly rated small institutions, for which the interval could be extended to 18 months." FDIC Report, *supra* note 8, at __.

[66] *See, e.g.*, Henderson and Tung, *supra* note 8, at __.

[67] The primary mechanism for examiner action is the bank's CAMELS rating. Examiners rate banks on a scale of 1 (good) to 5 (bad) in each of six areas – Capital adequacy, Asset quality, Management, Earnings, Liquidity, and Sensitivity to market risk – and then assign a composite score. A score of 1 means a bank is performing far above average; 2, the most common score, means "fundamentally sound"; 3 means "some degree of supervisory concern"; 4 means generally unsafe and unsound conditions; and 5 means severe problems and likely failure within one year. *See* OTS Examination Handbook, Section 070, pages 070A.3 & .4.

[68] This written commitment commonly comes in the form of a board resolution creating a Memorandum of Understanding between the bank and the regulator. *See* FDIC Report, *supra* note 8, at ___.

threat of legal enforcement, including orders to curtail particular activities. For instance, section 8 of the Federal Deposit Insurance Act (FDIA) gives various bank regulators vast power to curtail unsafe or unsound banking practices (that is, excessively risky activities) through the use or threat of cease and desist orders, removal of managers, or imposition of civil fines.[69] The threat of formal regulatory action is used to pressure banks into informal accommodation of regulatory wishes, through agreements known as "memoranda of understanding." When informal action is ineffective at restricting banking excesses, regulators have the power to take formal, legal action. For instance, cease and desist orders are authorized where bank regulators believe a bank's activities are likely to result in a "significant dissipation of assets or earnings" or are likely to "weaken" the bank or "prejudice" its customers.[70] Although rarely used, as we discuss and criticize below, these powers are construed to be broad enough to prevent excessive risk taking by banks. The government report on the failure of the bank Washington Mutual describes the examiner role this way: "[The regulator] is responsible for monitoring an institution's risk to the [the taxpayers]. [The regulator] had authority to perform its own examination of WaMu and impose enforcement action to protect the [the taxpayers]."[71]

The on-site examination process is meant to take place after loans happen but before it is too late. The examination process is designed to "identify the risk of failure in troubled institutions in sufficient time for supervisors to take corrective action" and therefore avoid social losses that cannot be remedied through traditional ex post litigation processes.[72] Another important tool available to examiners is a follow-up enforcement action, which is designed to "control the risk-taking behavior of problem banks after they have been identified."[73] Thus, banking regulation is not so much about preapproval for a decision to make a particular loan or type of loan, but rather an ongoing check on whether the lending decisions that have been made are acceptable from a social welfare standpoint. If they are not, the examiners have the power to force the bank to change its underwriting policies or raise additional money (that is, buy insurance against failure).

The regulatory veto is rare in law, but, as discussed below, we believe it is the essential feature of banking regulation. Why

---

[69] *See* 12 USC § 1818 (b, e, i(2), etc.)
[70] *See, e.g.*, 12 USC § 1818(c)(1).
[71] Wamu Report, *supra* note 7, at 35.
[72] FDIC Report, *supra* note 8, at 439.

give bank regulators the power to control business activities and even shut down entire firms? In the model of regulation we develop, the regulatory veto may be necessary if other forms of regulation are insufficient to optimize the amount of firm activity because of ex ante information deficiencies. We consider the various regulatory options, including no regulation, ex-post litigation, and the banking analog of command-and-control regulation, showing why they are insufficient to maximize the value of banking. We then show why adding the regulatory veto option can improve social welfare, but point out how it is commonly used in a suboptimal way that perversely undercuts its purpose.

C. *Regulatory Reform Proposals*

Commentators have offered a variety of explanations for the failure of bank regulation, having to do with issues such as the incentives of bank CEOs, banks being "too big to fail", and the failure of bankers, customers, and regulators to understand new financial products. Each of these alleged causes have generated reform proposals targeted at a specific alleged cause. Crucially, all of them fall into either the ex ante or ex post methods of regulation. In the ex ante category are restrictions on executive pay to change banker incentives to engage in particular activities, limitations on the size of banks (including the so-called Volcker Rule), new disclosures required for specific products, and a proposed requirement for government preapproval of financial products. In the ex post category are new rules requiring claw backs for certain monies paid to executives before accounting restatements, increased insurance requirements (in the form of convertible capital, for instance), and calls by many for more civil and even criminal cases against banks for the losses suffered by borrowers and taxpayers.

A series of recent papers focuses on ex ante incentives *of bankers*. Sanjai Bhagat and Roberto Romano propose paying bank CEOs with more long term equity, specifically that all equity options vest only after the CEO has left the bank.[74] Lucian Bebchuk and Holger Spamann propose changing the mix of CEO pay to include more debt in order to give bank CEOs better down-side risk incentives.[75] They want banks to do less of certain types

---

[74] Sanjai Bhagat & Roberto Romano, *Reforming Executive Compensation: Focusing and Committing to the Long-Term*, 26 YALE J. ON REG. 359 (2009). This proposal is similar to one made more generally by the authors. See M. Todd Henderson & James Spindler, *Corporate Heroin: A Defense of Perks, Executive Loans, and Conspicuous Consumption*, 93 GEO. L. J. 1835 (2005).

[75] Posner & Wyle, *supra* note 20, at __.

of activities, and propose achieving this by changing the rewards CEOs get from engaging in them.[76] In a friendly amendment, Fred Tung has proposed paying bankers in part with the publicly held subordinated debt securities of individual banks.[77] His argument is that the Bebchuk and Spamann proposal, which pays in the debt of bank parents (known as "bank holding companies"), would offer a noisy signal of the expected downside of particular risk taking.[78] He argues that paying in subordinated debt issued by individual banks, rather than their BHC parents, would offer better ex ante incentives for optimal risk taking.[79] These proposals are supported by empirical research done by Tung, along with co-author Xue Wang, finding that bank CEOs with greater amounts of compensation contingent of bank solvency (e.g., deferred compensation) took less risk than other bank CEOs during the run up to the Great Financial Crisis.[80]

A different set of papers focuses on the incentives *of regulators*. Fred Tung and one of the authors propose paying bank examiners with a mix of the equity and debt of the banks they regulate in order to improve their incentives to take actions to optimize bank regulation and intervene where necessary to limit bank losses.[81] In a follow-on paper, they propose improving examiner incentives further by using an auction method to allocate regulatory resources.[82] Under their proposal, examiners would choose banks to regulate rather than the other way around, thereby helping to ensure a better fit, the production of more information about bank risk, and to discourage regulatory capture. These proposals are the most related to the argument we present in this paper, since they focus not on ex ante or ex post regulation of banks, but instead on the incentives of regulators. As discussed below, our argument identifies regulatory failure as the primary source of recent banking failures, and highlights the need to improve regulatory design.

Focusing instead on ex ante restrictions on bank size, law professor Jonathan Macey and banker James Holdcroft, Jr. propose limiting the size of banks in order to reduce the costs of

---

76

77 *See* Frederick Tung, *Pay for Banker Performance: Structuring Executive Compensation for Risk Regulation*, 105 N.W. L. REV. 105 (2011).

78

79

80 Frederick Tung & Xue Wang, "Bank CEOs, Inside Debt Compensation, and The Global Financial Crisis, ____ (2011).

81 *See* Henderson & Tung, *supra* note 8, at __.

82 *See* M. Todd Henderson & Frederick Tung, *Reverse Regulatory Arbitrage: An Auction Approach to Regulatory Assignments*, *98* IOWA L. REV. 1895 (2013).

individual bank failure.[83] They argue that the only way the government can credibly commit to not bailing out banks that take excessive risks is to break them up so that failure is an option.[84] Their proposal would prohibit any bank from amassing liabilities that exceed five percent of the value of the FDIC insurance fund.[85] According to James Kwak, co-author of *13 Bankers: The Wall Street Takeover and the Next Financial Meltdown*, as of 2010, this would limit the size of any bank's liabilities to $3 billion, meaning the proposal would break up "over two hundred" banks.[86] (The asset-limiting proposal in *13 Bankers* is timid in comparison, proposing to break up only the six banks larger than $270 billion.[87]) The chief virtue of the Macey and Holdcroft proposal is that it takes advantage of the private information held by bankers about their efficiency at engaging in certain activities. They argue that their approach "does not require any restrictions on activities of banks or on the location of those activities of any kind." At the same, their one-size-fits-all proposal would result in "lost economies of scale" but they believe these would be "offset by the . . . savings realized by avoiding future bailouts."[88]

A related proposal—the so-called Volcker Rule—would limit banks from using their own cash to engage in certain types of investments to discourage risk taking. The rule would apply to all banks, regardless of their financial position, skill, history, or any other factor relevant to their ability to efficiently engage in this type of activity. In other words, these proposals apply a one-size-fits-all limit on the amount of activity any bank can engage in, regardless of its quality.

Even more aggressively, Eric Posner and Glen Wyle propose an FDA for financial products to pre-screen financial products along the dimension of social welfare.[89] This proposal envisions

---

[83] See Macey & Holdcroft, *supra* note 22, at __.

[84]

[85]

[86] *See* JAMES KWAK, 13 BANKERS: THE WALL STREET TAKEOVER AND THE NEXT FINANCIAL MELTODWN (2011); *see also* http://baselinescenario.com/2011/04/26/3-billion-banks/. Their proposed limit is 5 percent of the FDIC Deposit Insurance Fund, which itself is 1.15 percent of total insured deposits, so the limit would work out to $3 billion as of 2010.

[87] See KWAK, *13 Bankers*, *supra* note 86, at ___.

[88] Macey & Holdcroft, *supra* note 22, at __.

[89] *Id.*

government agents assessing the value and potential cost of bank activities before the bank engages in them.

Each of these proposals (except Henderson & Tung's focusing on regulators) shares three things in common. First, they are all attempts to reduce bank activity levels, either in general or in specific types of activities. Second, all of these proposed reforms amount to command-and-control regulation designed to substitute the judgment of "experts" in government for the judgment of the managers, shareholders, creditors, and customers of banks. Finally, all involve regulatory decisions being made with regard to neither a bank's profit opportunities nor regulators' information about systemic risks created by bank activities.

The reform proposals discussed above have little promise for limiting systemic risk. Some of them simply do not address the root causes excessive risk taking, and several of those that do are likely to be overly burdensome on productive banking activity and the economy in general. In other words, the reforms will either not work or err too far on the side of constraining bank activities at the expense of social welfare.

* * *

As we show in the next section, these modes of regulation have problems. All of ex ante regulation is faced with either a problem of ineffectiveness or limited information and overbroadness. Ex post measures are hampered by judgment-proofness and uninsurability due to systemic risk. The regulatory veto, while promising as a means of increasing the information available to regulators and allowing bank-by-bank application of prudential rules, suffers from a serious structural problem that, in its current form, renders it ineffective.

III. A SIMPLE MODEL OF REGULATION OF BANK RISK-TAKING

In this part, we offer a simple model examine the various regulatory choices available to regulators, and to explore the strengths and weaknesses of each. As it turns out, there are serious shortcomings of ex ante and ex post regulation, as well as the regulatory veto.

Ex post measures of risk regulation (namely, post-harm litigation and mandatory pre-harm insurance) will not work

because it is in precisely those cases where risk materializes as reality that the firm will be insolvent and unable to pay. Insurance, a common solution to the judgment-proofness problem, is ineffective because systemic risks are likely to put unbearable pressure on insurance companies.

Ex ante regulation avoid the problem of judgment proof banks. However, all forms of ex ante regulation of banking activity suffer from a severe informational problem: the regulator acting ex ante does so based on limited information in prescribing what actions may be taken, and does not take into account information it may subsequently learn about the firm's production and potential systemic risk. In addition, ex ante regulation is typically of a one-size-fits-all variety, and fails to allow for heterogeneity among firms, financial products, and customer needs. As a result, ex ante regulation of banking activity is unlikely to yield the optimal level of banking activity. It will either prove ineffective and gameable by industrious banks, or else be a significant drain on both the banking sector and the wider economy. These problems are generally true of all the forms of ex ante regulation we identify: size limits, prudential regulation, financial transactions taxes, command and control, and incentive based approaches.

In addition, as we show, some forms of ex ante regulation are poorly designed to limit systemic risk. For instance, prudential regulation, such as the capital adequacy ratios required under the Basel Accords, does little to prevent systemic risk-taking. Prudential regulation may serve a purpose in protecting the financial firm's creditors: it constrains borrowing (and hence activity levels) relative to the amount of equity shareholders have at risk in the firm. By construction, however, prudential regulation based only on debt/equity ratios does not guard against systemic risk or externalities outside of the firm's direct stakeholders, and hence cannot generate optimal social welfare incentives.

We then turn to what we call the "regulatory veto" – the system of bank examination and the bank examiner's power to shut down a bank based on examination results. Unlike the other forms of ex post regulation, litigation and insurance, the regulatory veto is not subject to the problem of judgment proofness because it does not wait for the harm to actually occur. Unlike ex ante regulation, the regulator does not prescribe limits on banking activities, and is able to wait and consider additional information before making a decision on letting the bank operate or shutting it down before greater harm is incurred. In other words, ex ante regulation happens too soon, while ex post regulation happens too late. The regulatory veto could be, as Goldilocks said, just right.

Unfortunately, as we then show, the regulatory veto has a serious structural problem: by allowing the bank to move first, the

regulator finds itself subject to a form of the ultimatum game, and hence may accept banking activities that are, on net, harmful to society. Hence, the regulatory veto will require some fixes to work well, a problem we turn to in Part IV.

A. The Model

To explore how banking regulation works – or, as it largely turns out, does not work – we utilize a simple economic model. Our model economy consists of (i) a bank (or, equivalently, a financial firm engaged in shadow banking), (ii) the bank's shareholder, who manages the firm and earn its profits, (iii) the bank's creditors, who provide much of its working capital, and (iv) the regulator, who may undertake various sorts regulatory measures to attempt to maximize social welfare. The crux of the model is that equity holders (and, so some extent, the firm's bondholders) have an interest in taking more risk than is optimal from a societal standpoint, owing to the limited downside of both debt and equity holders.

One criticism of the model, which we point out here at the beginning, is that there is no separate manager; in contrast, this is a banking firm without agency cost, run completely in line with the shareholders' interests. We omit a measure of agency cost for two reasons. The first is to demonstrate how pernicious the problem of bank shareholder incentives are with regard to excessive risk taking: even without assuming bad bank executives (a common refrain heard in the wake of the Financial Crisis), we show that exactly the same excessive risk taking behavior is to be expected. The second reason to omit agency cost is that it significanly complicates the analysis – among other things, the manager can sometimes be turned usefully against the sharheolders – and is deserving of its own treatment, which we do in a separate paper [or papers] in which we examine various proposals to regulate bank manager pay.

1. The banking firm and its projects

Our bank or financial firm is funded with some amount D of debt and an amount e of equity. The firm uses this capital to invest in financial assets. For simplicity and concreteness, we will suppose that each project costs a dollar. However, banking projects vary in the following ways. First, there are two general types of banking projects, low and high risk projects, denoted as l and h projects, respectively. These projects are essentially coin flips of varying amounts. We will assume that each of the l and h projects has a binary distribution, and that all of the projects are perfectly correlated with one another.[90] Only a certain number of each type of project is profitable: the first n_l of the low risk projects pay off either $\$2+\pi_l$ or π_l in the case of success or failure, respectively, with the probability of success/failure at 50%. The remainder of the low risk projects pay off only \$2 or \$0, so that their expected profitability is zero.

Similarly, the first n_h high risk projects pay off $\$3+\pi_h$ or $-\$1+\pi_h$ for success/failure with probability 50%. These first n_h projects have an expected value of π_h, while all the other high risk projects have an expected value of zero.

Both the number of projects that are profitable (n_l and n_h) and their degree of profitability (π_l and π_h) are random variables whose distribution is common knowledge. The firm and its shareholder observe the realization of these variables prior to undertaking any banking activity. The regulator, in contrast, knows only the distributions of these variables ex ante, though over time the regulator may gain more information about their realizations. In short, it is socially beneficial for the bank to undertake a certain number of banking projects; beyond that, however, all the the bank does is increase risk, both for its security holders and, as we discuss below, for the society in general.

2. Shareholders

There is a single unity shareholder who both owns and runs the bank. The shareholder has some private wealth endowment of

---

[90] While perfect correlation across all projects is certainly an unrealistic assumption, it qualitatively changes little in our analysis. The addition of more projects, whether perfectly correlated or independent, always increases the bank's overall riskiness. In a prior draft, we assumed imperfectly correlated normal distributions for bank projects; results there differ little from here.

 which she may invest in either banking or non-banking projects. Non-banking projects pay a market rate of return of *r*, while banking endeavors are allowed to vary according to the level and type of investment.

Payoffs to the sharheolder are not the same as the firm's expected value, owing to the option character of equity: in a firm with limited liability, equity holders effectively have the option to purchase the firm's realized cash flows for the value of the oustanding debt. Notably, shareholders are free to walk away from a firm that creates even catastrophic financial injury to others.

3. Bank creditors

The bank's capital structure also includes creditors, who have limited upside and no control rights over the actions of the banks. Creditors, instead, can only discipline the bank by withholding their credit from it. Depositors, for instance, who believe a bank is poorly managed or undercapitalized, such that depositors fear for their deposits, will rationally run on the bank, causing it to shut down. Bondholders, similarly, will refuse to lend new funds to the bank, and refuse to roll over term debts as they come due. Counterparties (those who engage in transactions such as securities repurchase agreements and propriety trading with the bank) may also choose to stop doing business with the bank, robbing it of a source of short term funds.

Unfortunately, the monitoring abilities or incentives of these creditors may be severely limited. Depositors, who are covered by FDIC insurance, have no incentives to run on a poorly run or undercapitalized bank. Bondholders lend for a term to the bank, and hence have limited ability to get their funds back as conditions change. More controversially, perhaps, the counterparties of a bank may not be able to monitor the bank's creditworthiness due to the complexities of the trades or the anonymity of the short term money markets.[91]

What all this means is that the bank's cost of capital, at least in the short term, will not reflect the actual risks that the bank imposes on its creditors. Depositors simply do not care; counterparties do not always know with whom they are dealing; and bondholders have limited ability to act in the short term. Banks and their shareholders, then, have the ability to externalize much of their costs from risky activities onto their creditors.

---

[91] *See, e.g.*, Viral Acharya and Alberto Bisin, Counterpary Risk Externality: Centralized Versus Over-The-Counter Markets, NBER Working Paper 17000, *available at* www.nber.org/papers/w17000.

4. The regulator and systemic risk

Banking activity is governed in some way by a banking regulator who acts to maximize total societal welfare.[92] Bank regulation is required in order to limit the externalities that the bank and its shareholder can impose on others. Ordinarily, of course, one would think that externalities imposed on creditors require no government intervention: the firm's cost of capital simply rises, as creditors with rational expectations demand a higher return of banks that take on additional risks or undesirable projects, or, alternatively, they require some monitoring and control rights, such as a seat on the board, to manage the firm concordantly with their interests. However, as discussed above,[93] these creditors do not necessarily have the ability to fend for themselves: the FDIC does not get to choose where depositors put their deposits, counterparties may not be able to sufficiently analyze their trades and the bank's creditworthiness,[94] and even bondholders may be subject to short term expropriation.

Additionally, and perhaps more importantly, is the concept of systemic risk. The failure of a bank may lead to consequences far beyond just the bank's various securityholders. Banks are always lending to one another in the money markets, such that the failure of one bank imposes losses on another. Banks often invest in the same sorts of assets, such as residential mortgages. When a bank fails and attempts to sell its assets, prices may be pushed lower in a form of "fire sale," impacting other banks' balance sheets and forcing them to sell assets as well.[95] This is sometimes referred to as financial contagion; though the existence and mechanics of financial contagion and fire sales are controversial issues, the conventional wisdom is currently that they exist and do impose significant losses on society.[96] In any event, problems at

---

[92] Although regulators may be motivated by private interests as well, see, e.g., M. Todd Henderson, "The Changing Demand for Insider Trading Law," Working Paper 2012, we assume they are motivated by public interest.

[93]

[94] *See, e.g.*, Nils Beier, Holger Harreis, Thomas Poppensieker, Dirk Sojka, & Maria Thaten, Getting to Grips with Counterparty Risk, McKinsey Working Papers on Risk, Number 20, June 2010 (available at http://www.mckinsey.com/client_service/risk/latest_thinking/working_papers_on_risk)

[95] *See, e.g.*, Douglas W. Diamond and Raghuram G. Rajan, *Fear of Fire Sales, Illiquidity Seeking, and Credit Freezes*, The Quarterly Journal of Economics, Vol 126, May 2011 (pp. 557 – 591).

[96]

one bank may implicate or even infect other banks, leading the whole sector downward.

Finally, banks play integral roles in the modern economy. Producers rely on banks for working credit and to finance capital investments. Homeowners (and the real estate industry) depend upon banks finance home building and sales. Consumers rely on banks for payments systems. And the entire economy relies upon bank lending to maintain the money supply. The failure of a significant portion of the banking industry, therefore, significantly harms the wider economy, as it appears to have done in the recent Financial Crisis.[97] The fact that such significant externalities exist, and are so widespread in scope so as to affect the entire national (and, potentially, international, economy), implies that a regulatory solution at the national level is desirable.

In our model, then, we consider the bank's systemic risk that it externalizes onto the wider economy and allow the regulator to take this systemic risk into account when deciding upon regulatory strategy. Specifically, we assume that the amount of systemic risk (which is, for our purposes, the expected harm imposed on society by the bank), which we denote as a function of the number of low and high risk projects undertaken, $R(l,h)$, is increasing in the actual portfolio risk of the bank, and at an increasing rate (formally, $R' \geq 0$, $R'' \geq 0$). That is, the marginal effect of each project on overall systemic risk is increasing. Adding a project to a bank with few projects has little effect on systemic risk, since the bank is relatively quite solvent in the event that the project fails. In contrast, adding an additional project to an already heavily invested bank may well be the straw that breaks the camel's back, and may go further by causing losses at other banks and in the wider economy. One possibility, which we discuss later on,[98] is the possibility that activities at one bank affect the systemic risk created by activities at another bank; for instance, an additional investment in sub-prime mortgages creates more risk given that other banks are already heavily invested in the area.

With this in mind, our regulator in the model may undertake any of several forms of regulation with the goal of maximizing expected overall social welfare. This includes the gains and losses of the shareholder, creditors, and the wider society (in the form of

---

[97] It is not universally accepted that the Financial Crisis caused the ensuing recession. For example, Tyler Cowen argues that reductions in productivity is a key factor that explains the Great Recession. *See* TYLER COWEN, THE GREAT STAGNATION (2011); see also Jeffrey Miron & Natalia Rigol, "Bank failures and output during the Great Depression," *http://www.nber.org/papers/w19418*.

[98]

systemic risk, *R*). We can make this problem somewhat simpler, as given the creditors' investment, creditors' expected losses are shareholders' expected gains given the expropriative nature of risk-taking in this model. Hence, the regulator seeks to maximize the sum of the bank's expected profits less the systemic risk that bank imposes on society.

We consider three classes of possible regulatory actions: ex post litigation and insurance covergage, ex ante direct activity level restrictions, and ex post intervention to shut down the bank's activities (at a cost) after shareholders and bank managers have made their choices. An additional class of regulation – mandated executive compensation characteristics – we leave to a future project, as it requires a consideration of the more complicated agency cost problem.

B. How do the various alternatives fare?

In this section we consider the operation of our model firm and economy under several modes of regulation, drawing from both the law and economics literature and current modes of bank governnace.

1. A baseline: laissez-faire

As a baseline, consider the laissez-faire case in which regulators do nothing, allowing private individuals or firms to act in their own interest. This approach has some appeal, relying, as it does, on private market participants to act on their own information according to their preferences. Shareholders, managers, creditors, and counterparties make their own investment decisions to maximize private gains.

The effectiveness of this laissez faire approach depends heavily, however, on whether there are any potential costs generated by the activity that are not fully internalized by the individual or firm, and therefore not priced by the market. Without an internalizing mechanism and in the presence of Coasian transactions costs,[99] negative externality generating products and activities will receive too much investment.[100] A polluting factory is

---

[99] A potential market solution, following Coase, is for those injured by the externality to contract with the injurer to refrain from the harmful conduct. *See* Ronald H. Coase, *The Problem of Social Cost*, 3 J. LAW & ECON. 1 (1960). In general, where benefits or harms are dispersed, the large costs associated with collective action and contracting make market solutions of this sort unlikely. Further, such solutions typically rely on enforceable contracts.

[100] Pollution is the classic example of an externality that cannot be internalized without regulation of some kind. If not charged for emissions that create acid

the prototypical example, since a factory that can force local farmers to bear some of the costs of production will overproduce relative to the social optimal. In this way, excessive bank activity leading to systemic risks can be thought of as a sort of financial pollution. And just as with unregulated emissions from factories and power plants, there is abundant evidence of financial pollution.

Given any level of debt investment in the bank's capital, shareholder will tend to choose a higher level of banking activity than is optimal. Consider a specific case where the bank has FDIC-insured deposits of $10, the shareholder has an endowment of $2, and the market rate of return is 0%. There are 10 each of high risk and low risk projects available to the bank. Further, the shareholder knows that the first 3 low and high risk projects carry with them an expected profit of $0.20 (i.e., $n_l = n_h = 3$, $\pi_l = \pi_h$ =$0.20), and each low risk project increases expected social costs (in terms of systemic risk) by $0.08, while each high risk project increases expected social costs by $0.16 (i.e., $\frac{\partial R}{\partial l}$ = $0.08, $\frac{\partial R}{\partial h}$ = $0.16). In such a case, the social optimum is attained by undertaking all 6 of the profitable projects, which yields aggregate returns to the bank's shareholders and debtholders of $1.20, and expected costs due to systemic risk of $0.72, for an overal social return of $0.58.

Unfortunately, absent regulation of some sort, the shareholder will not undertake the socially optimal set of projects. In the event that the projects fail (and recall that, by assumption, they either all succeed or all fail), the shareholder winds up with zero, since the assets of the bank will be insufficient to fully reimburse the creditors (here, the FDIC).[101] This means that the shareholder enjoys the upside on risky projects, but is indifferent to the degree of failure that the bank may experience. So, for the first three low and high risk projects, the shareholder's expected payoff is 50% * ($2 + $0.20) = $1.10 and 50% * ($3 + $0.20) = $1.60 on each project, respectively. For all other projects such that $l + h \leq 10$, the shareholder's payoff is $1 and $1.50 for low

---

rain that destroys the crops of farmers in distant locals, industrial plants will produce too much.

[101] One can verify this by noting that in the event of failure, for any number of projects undertaken that is greater than the equity investment (i.e., $h + l > e$), the revenues will be insufficient to repay the full $10 of debt. And, clearly, the shareholder would, in these circumstances, always choose to invest the full amount of debt capital – the upside is essentially free to the shareholder in this case; the choice of how much equity to contribute, in contrast, depends on the project parameters.

and high risk projects, respectively. The hierarchy of investment, then, is the profitable high risk projects (expected gross payoffs to the shareholder of $1.60), the non-profitable high risk projects ($1.50), the profitable low risk projects ($1.10), and, finally, the the non-profitable low risk projects ($1). The shareholder will undertake all 10 high risk projects and two of the three low risk profitable projects. The shareholder's expected net payoff from this is 3 * $1.60 + 7 * $1.50 + 2 * $1.10 - $10 - $2 = $5.50. Overall social welfare is, however, negative, at 3 * $0.20 + 2 * $0.20 – 10 * $0.16 – 2 * $0.08 = -$0.76.

Two things are apparent from this example. First, the shareholder will engage in too much risk overall: even if all the profitable projects are used up, there is an incentive to simply maximize the option value of equity's limited liability. Second, because of the shareholder's benefit from risk, the shareholder may even choose to pass up profitable projects in order to engage in unprofitable risk-seeking strategies.

2. Ex post: litigation and insurance

Because banking is widely thought to generate significant negative externalities, ex post litigation—analogous to tort causes of action in the pollution context—is available to force parties who act in socially costly ways to bear the costs of doing so. After the harm has been done, injured parties may sue to recover their losses from the offending bank.[102]

Ex post litigation has the virtue of keeping regulatory interference to a minimum and avoiding decisions based on limited information, and is a common regulatory option for that reason. For instance, to evaluate the social welfare of the design of nearly all consumer products, we use ex post litigation (that is, tort law) instead of preapproval by government bureaucrats. This choice reflects the fact that the expected costs of preapproval (that is, decision costs plus error costs) are extremely high in light of the lack of information held by regulators ex ante. Regulators do not know, for instance, the products or features consumers will demand, the products producers will be able to design and manufacture efficiently, or the social costs of products before they are made and used. In the banking context, then, while a regulator may suspect, but not know for sure, that a financial firm's

---

[102] For instance, in August 2011, Bank of America agreed to pay $8.5 billion to settle a suit brought by disappointed mortgage investors. See, Nelson Schwartz, *Bank of America Settlement Faces Growing Challenges*, N.Y. TIMES DEALBOOK, Aug. 30, 2011, *available at* http://dealbook.nytimes.com/2011/08/30/homeowners-seek-to-block-bank-of-america-settlement/.

activities impose a social cost of, say, $10 approximately half the time, absent other concerns, it would be preferable for the regulator to simply wait to see whether the harm actually materializes, and of what magnitude it is, before assigning liability.

This is true in other areas of financial regulation as well. Take securities law. For many years, state laws known as "Blue Sky" laws authorized state securities regulators to prescreen securities issued by private firms to see whether or not they were reasonable investments for citizens of the particular state.[103] The information problem of ex ante regulation is plain, since assessments about the riskiness of a security or a particular firm are famously difficult to make. The recent experiences with the Internet stock bubble and the credit derivative bubble of that lead to the Financial Crisis are ready examples of the errors that can be made in trying to evaluate values and risk ex ante. It was perhaps this problem that led to the New Deal Congress that passed the federal securities laws to reject so-called "merit" regulation, relying on ex post litigation (primarily on fraud and disclosure-related causes of action) to give issuers of securities proper incentives.[104] In addition, state merit regulation has largely died,[105] as the costs of pre-screening (that is, decision costs plus error costs) are thought to be unnecessary given the private incentives and availability of ex post sanctions.

There are two major shortcomings of ex post liability. First, if the injured parties are diffuse and have relatively small stakes, then the costs of bringing a lawsuit may be prohibitive. This is likely true in cases of environmental harm, where many people may be injured slightly by pollution, and in banking, where a bank failure may impact thousands of depositors. There are mechanisms for coordinating litigation and reducing these costs, including class action litigation and government-brought cases. But it may be more efficient to rely on ex ante regulation, such as limits on emissions, cap and trade systems, or bans, depending on the economic, political, and practical circumstances of the particular case.

---

[103] *See, e.g.*, Jonathan Macey & Geoffrey P. Miller, *Origin of the Blue Sky Laws*, 70 TEXAS LAW REVIEW 347 (1991) (describing Blue Sky laws, and offering a public choice explanation for their passage).

[104] For a discussion of the regulatory choices and environment surrounding the passage of the Securities Act of 1933 and the Securities Exchange Act of 1934, see JOEL SELIGMAN, THE TRANSFORMATION OF WALL STREET (3d Ed.) (2013).

[105] The National Securities Market Improvement Act of 1996 amended section 18 of the Securities Act of 1933 to preempt state blue-sky laws regarding a specific class of "covered securities," including securities traded on national exchanges. *See* 15 U.S.C. §77a *et seq*.

Second, the injurer must be able to pay any judgment in order to internalize the social costs of the activity. Where an actor is likely to be bankrupt when the injury occurs, or be bankrupted by the injury, the actor will not expect to bear the full costs of the injury, and will therefore engage in a supra-optimal level of the injuring activity. The problem is often more acute when the harming party is a firm: shareholders, who are protected by limited liability, may prefer projects with high insolvency risks, since the shareholder has, effectively, an option to purchase the firm's cash flows for the price of the outstanding debt. Hence, in industries where the risks are great and losses are often very large, as in pharmaceuticals and banking, we expect tort liability to be displaced or complemented by other regulation.

Judgment-proofness is clearly a problem in our banking model. The firm will be insolvent and unable to pay creditors and injured third parties in the event of severe financial injury to others. While presumably the FDIC would be able to sue for its losses, and the federal government would similarly be able to attach assets to cover bailout or other such taxpayer costs, this will generally happen too late, after the bank has gone bust. In the numerical example above, for instance, the bank undertakes 10 high risk and 2 low risk projects. In the event that the bank goes bust, the bank's net assets will be the 5 * $0.20 = $1 produced by the five profitable projects (3 high and 2 low risk) less the 10 * -$1 = -$10 losses caused by the failure of the 10 high risk projects, for a net value of -$9. Absent the problem of FDIC insurance, this would not be all too concerning, as creditors would have bargained for either controls or an interest rate that gives them an expectation of at least breaking even. In addition, however, the bank will have generated $1.76 in expected costs due to increased systemic risk – which, when realized, could be much higher. The bank, now insolvent, is clearly in no position to repay either its tort or contractual creditors.

One common regulatory response to the judgment-proofness problem is to require insurance. The regulator can mandate that the potential bad actor pay premiums to a regulated intermediary who in return guarantees the actor's specified liabilities.[106] Another option is to require bonding by the firm. The regulator can require the actor to set aside money or assets in a segregated account that can be attached in the event of subsequent injury to others.[107]

---

[106] The regulator need not be the government, since a self-regulatory organization may do this by using a centralized third-party clearing house to achieve the same result.

[107] The leading treatment of this issue is found in Steve Shavel, *Minimum Asset Requirements and Compulsory Liability Insurance as Solutions to the Judgment-Proof Problem*, 36 RAND J. ECON. 63 (2005).

Either measure amounts to a pre-commitment by the insured or bonder to pay future judgments, at least up to a specified amount. Bonding imposes costs directly on the actor, while insurance requires oversight by the insurer to promote optimal behavior; the insurance actuarial process ideally takes into account bank activities in setting premiums and extending coverage, such that banks have an ex ante incentive to promote social well-being.[108]

Unfortunately, in the banking context, insurance may be ineffective due to the correlation of failures among firms.[109] Indeed, precisely because it is systemic risk that provokes concern about bank collapses, private insurers are unlikely to be able to provide meaningful protection. Insurance works by aggregating risks across the insured, so that the costs incurred by one party can be dispersed across the entire pool of insureds. But if financial firms engage in similar projects with the same risk profiles, then all firms will face insolvency risk at the same time. We have assumed in our model perfect correlation among projects, so that all firms will be insolvent at the same time.[110] In such a case, insurance will not have any impact on firm risk taking. While the perfect correlation assumption is clearly unrealistic, we still reach the qualitatively same result with regard to insurance so long as a significant degree of correlation exists among the projects and so long as the potential injuries are large. Under more realistic assumptions in the model, when one financial firm experiences failure due to systemic risk, so will many others; insuring systemic events such as this is often beyond the scope of private insurers.

The intuition behind this result is born out to some extent by recent experience with the failure of financial firms. While insurance for default risk in banking does exist in the form of, among other things, credit default swaps, systemic events may lead to industry-wide (and economy-wide) failures and catastrophic loss amounts. This happened, for instance, with AIG credit default swaps written as protection on mortgage-backed securities; not only did certain bank assets fail, but much of AIG's asset portfolio

---

[108] FDIC insurance is only weakly linked to risk because of the danger of self-fulfilling prophecy. Risk-based premiums are provided in 12 U.S.C. § 1817(b)(1)(A), (C), and regulations are set forth in 12 C.F.R. §§ 327.09-327.10.

[109] It is not clear that even the FDIC could adequately cover a single large bank failure. In the Washington Mutual failure, for instance, the FDIC arguably rushed Washington Mutual into receivership due to concern over the ability to fully insure its deposits in an abysmal failure. *See* Kirsten Grind, The Lost Bank (interview describing FDIC concerns available at http://failuremag.com/feature/article/the_failure_of_wamu/)

[110] Perfect correlation means that firm risk increases linearly in the number of projects; independence would mean that firm risk increases in the square root of the number of projects.

declined in value as well.[111] Insurance in this instance was woefully ineffective to properly incentivize firms, and the common narrative is now that insurers themselves expected strategic judgment-proofness to work in their favor.[112]

3. Ex ante regulation: size and asset type restrictions

Given the failures—both theoretical and, apparently, real—of both litigation and insurance in governing bank behavior, some additional method of regulation is necessary. We turn our attention now to several variations on what we generically term ex ante bank regulation – activity restrictions of various sorts – which constrain either the amount of banking activity that a firm may undertake, or else the sorts of projects available to the firm.

Analogizing to other industries, activity restrictions are similar command-and-control regulation, where experienced bureacrats prescribe mandatory best practices for the firm. Command-and-control is most commonly found in areas of relatively slow development of products and technologies, and where informational asymmetries between regulators and firms are low. For example, in stable industries, like energy utilities, the costs and benefits of the activities are well known, and the amount of innovation is relatively small (putting aside the concern that low innovation results from the degree and form of regulation). Energy producers rarely produce new products, and regulators can reasonably estimate the future social costs and benefits of activities based on historical experience. Such judgments are made on the basis of rigorous cost-benefit analysis, as it is relatively straightforward to estimate with some scientific precision the impacts of regulations on individuals and firms. The EPA can therefore be relatively confident that by limiting firm activities in particular ways it is not causing underproduction (or allowing overproduction) of valuable products and services; the EPA might, for instance, mandate a cap on particulate emissions from a particular sort of factory, as well as prescribing production methods, effectively limiting production to a set specification,

---

[111] *See* Richard Squire, *Shareholder Opportunism in a World of Risky Debt*, 123 HARV. L. REV. 1151 (2010). While Squire writes that AIG would not necessarily have been insolvent based solely on the policies it wrote, its assets (largely placed in real estate) declined significantly in value at the time of the banking crisis.

[112]For instance, the International Association of Insurance Supervisors issued a position statement on the financial crisis noting that the insurance industry itself is susceptible to systemic risk. (available at http://www.iaisweb.org/__temp/IAIS_Position_Statement_on_Key_Financial_Stability_Issues.pdf).

based on extensive knowledge of the products created and the harms imposed by those emissions.

This ex ante regulatory approach is, however, of questionable efficacy in fast-evolving areas such as technology or structured finance, where the costs and benefits of new products are likely much more difficult to determine in advance. Finance also contains many dispersed heterogeneous innovators, producers, and consumers, again unlike centralized energy production. It is perhaps for this reason that cost-benefit analysis, at least as rigorously practiced by regulators like the EPA, has historically not been a part of the regulation of securities and other financial products.[113]

This points to a general drawback of activity level and other forms of command and control regulation: the level of overall social welfare is constrained by what the regulator knows at the time of creating the regulatory structure. As such, ex ante regulation fails to take account of valuable information held by the firms, by consumers, and by the regulator after observing production and usage for some time. This may mean that the activity level regulation inefficiently constrains production to a suboptimal level, as in the case where the regulator overestimates the costs or underestimates the benefits of the activity. Or, quite the opposite, this may mean that the activity level regulation is not strict enough.

An additional problem of being the first mover is regulation will necessarily have to be designed to anticipate behavior, and this means that behavior can be adjusted to avoid the regulation. Loopholes may be game-able by the firm; firms may simply switch from the proscribed means of production into more harmful or less efficient ones. When a firm's first-choice project becomes more costly due to regulation, the firm may switch to a second-choice project, which may be done with increased risk in order to boost returns to be similar to the first-best project.

The use of credit derivatives is an example of this sort of regulatory arbitrage. Capital adequacy rules require banks to hold cash in reserve, based on a risk calculation. To avoid holding cash,

---

[113] The D.C. Circuit recently reversed and remanded a SEC rule regarding shareholder access to the corporate proxy on grounds that the SEC did not engage in cost-benefit analysis of the rule. *See* Business Roundtable v. SEC, No. 10-1305, 2011 BL 191644 (D.C. Cir. July 22, 2011). Thereafter, the House passed a bill that would have mandated the SEC to do cost-benefit analysis for every rule. *See* Andrew Ackerman, *House Lawmakers Pass SEC Cost-Benefit Bill*, WALL ST. J., May 17, 2013, available at http://online.wsj.com/news/articles/SB10001424127887324767004578489410238066202. The bill went no further.

which after all does not earn a return, banks engaged in risk-transferring contracts known as credit derivatives. Banks would make loans, for mortgages, for instance, and then transfer some of the risk of the mortgage to other investors. This allowed them to hold less cash in reserve. For example, a loan of $100 would be packaged with other loans in a special-purpose entity, and then interests in the portfolio of loans were sold to investors. In this way, the bank could make new loans without holding cash in reserve for the original loans. The motive was to avoid the regulation, but the market grew because of an interest on the part of investors to participate in these deals without building a lending infrastructure or customer relationships. The result was a huge boom in the credit derivatives market, which eventually exceeded tens of trillions of dollars. When this market collapsed, the result far exceeded the original risk that was sought to be avoided by the capital regulation in the first instance.

Despite these general drawbacks, several recent proposals for bank regulation involve some form of pre-set activity limitation. These include proposals to limit the size of banks, restrict leverage or balance sheet growth, impose taxes on financial transactions, and limit the financial products that may be made available, as a form of financial "FDA." We consider these now in turn.

i. Firm size limits

A common proposal to constrain bank behavior is to simply limit the size of banks. For instance, economist and former Secretary of Labor Robert Reich has written that "the only way to make sure no bank is too big to fail is to make sure no bank is too big."[114] Several academics and banking observers have also proposed concrete limits on bank size.[115] These proposals are based on a conmmon sense view that if banks are constrained from growing their balance sheets, the level of risk undertaken by any individual bank will be smaller.

From the analysis of the shareholder's decision in the absence of regulation,[116] we know that shareholder will have the bank engage in too much activity in general. In our model, shareholders will generally prefer more banking activity than regulators; therefore, any limit set by the regulator will likely bind. In such a case, the regulator chooses activity level limit a, which is

---

[114] *See* Robert Reich, *Break up the banks: Why we must limit the size of banks*, CHRISTIAN SCIENCE MONITOR, April 6, 2010 (available at http://www.csmonitor.com/Business/Robert-Reich/2010/0406/Break-up-the-banks-Why-we-must-limit-the-size-of-banks, last visited July 11, 2012)

[115] *See* Macey & Holdcraft, , *supra* note 22, and accompanying text.

[116] *Supra* part [ ].

the maximum number of projects that a firm can undertake, with the expectation that the limit will be binding; the size of the firm's balance sheet will then be a.

An obvious problem is that the regulator does not observe the realization of either the number of profitable projects (n_l and n_h), or the level of profitability of those projects (π_l and π_h). So, for purposes of setting size limits, the regulator relies simply on the distribution of those variables, which we assume to be common knowledge. The greater the variance of these distributions, the less likely is the regulator's rule likely to be helpful, and the lower will be expected social welfare. The shareholder knows the realizations of these variables, but because of her conflicting incentives, simply telling the regulator these values is not credible. This basic problem – the limited knowledge of the regulator acting before activity is undertaken – is a general problem with ex ante bank regulation of all forms.

Consider also the bank's and shareholder's incentive given a size limit of a. Returning to our numerical example above, suppose that the regulator limits banks to a balance sheet size of \$6. This happens to correspond to the number of profitable projects available to the bank (recall that $n_l = n_h = 3$, $\pi_l = \pi_h$ =\$0.20). But this does not mean that the bank will invest in only profitable projects; to the contrary, the bank will choose to undertake 6 high risk projects, only 3 of which are profitable, leaving aside the 3 profitable low risk projects. What the size restriction has done, in this case, is simply to drive the bank into exclusively high risk projects. We think that this is likely to be a problem in the real world: banks who are constrained in the size of their balance sheets have incentives to get more bang for their buck, in the form of higher risk projects, even if those projects are not socially desirable.

There are yet other reasons to think that this approach is highly inefficient. Partly this is because limits on firm size are, by design, a one-size-fits-all solution. Our model assumes one representative bank; in reality, though, firms are, of course, different, and some firms may have many more profitable projects than others; management may be better, for instance, or the opportunities open to that bank may simply be of higher quality. In terms of our model, if we add more banking firms and allow n_l, n_h, π_l and π_h to vary across firms, the regulator's one size fits all rule is necessarily imperfect – even in the case of perfect information. Putting this in the context of the real world: it may well be that some banks are justified in being quite big, while other poorly run banks ought to remain quite small.

The nature of systemic risk, and how it evolves over time across the banking sector, presents another problem with ex ante regulation. The regulator may receive subsequent information about both the profitability of the firm and the systemic harm that the firm is likely to impose. This is especially true for systemic risk, where the likelihood of a systemic event depends upon what other banks and financial firms are doing, and is something that the regulator is uniquely situated to observe. For instance, in a two bank world, if Bank 1 is heavily invested in sub-prime real estate, but Bank 2 is not, Bank 1 likely carries far less systemic risk than in the counterfactual case where Bank 2 holds highly identical assets on its balance sheet.

This issue of similar asset holdings across banks suggests an additional shortcoming of size limits. If the sum of expected systemic harm across banks – R in our model – is a function of the sum of all projects undertaken, then the number of banks into which those projects are dividided is meaningless. The rationale of size limits assumes that projects abandoned by one bank (as it hits its asset limit) will not simply be picked up by additional banks – that is, there are a fixed number of banks, such that size limits really do limit overall banking activity. If new banks simply take on those projects, then the industry-aggregated bank balance sheet and portfolio correlation is unchanged. Suppose, for example, that in the absence of regulation, Bank 1 would undertake 100 projects, imposing an expected cost of $100,000 on society. If the regulator limits Bank 1 to only 50 projects, and no other banks arise or capitalize on the foregone projects, the regulator has successfully limited risk. If instead, however, Bank 2 is formed to snatch up the remaining 50 projects, there is no reason to think that the overall expected costs and benefits imposed on society will be any different than in the unregulated case with only Bank 1. They are, after all, the exact same projects, separated only by the corporate fiction. Unless there is something significant with regard to the particular corporate shell itself that houses the assets,[117] then size limits will do little to prevent opportunistic risk-taking. While a convincing account of systemic risk is required to definitively answer whether cabining projects in separate corporate shells alleviates systemic risk, no such account exists; until one is provided, we think size limits are unlikely to help.

---

[117] It is conceivable that there are significant real differences to having the same number of assets in one bank as opposed to spread homogeneously across two banks. For instance, Reich, *supra* note 114, suggests that larger banks have economies of scale in buying off politicians and regulators, suggesting that "too big to fail" is really a product of public choice.

ii. Prudential regulation and capital adequacy requirements

One of the primary ways in which banks are currently regulated is by setting a limit on the amount of banking activity that can be undertaken relative to the amount of shareholders' equity. Such principles are embodied in the capital adequacy ratios under the Basel Accords, discussed above.[118] Supposing, for instance, that the regulator mandated a capital adequacy ratio of 8 percent (as is required under Basel I), a bank with shareholder equity of $100 million would be allowed to engage in only up to $1.25 billion of financial activity.[119] If the bank or its shareholders wished to engage in more activity, they would have to raise more equity capital.

This type of prudential regulation is designed to protect the firm's creditors. A required equity cushion based on the amount of debt helps assure creditors that they do not take on more risk than initially contemplated. It also gives equity investors some downside, which may reduce risk taking to an extent. The problem, however, is that prudential regulation of this sort does little to protect those outside of the firm and to prevent large amounts of risk-taking activity.

Return to our numerical example from Part III.B.1. Suppose our shareholder has put $0.87 in the bank. Under an 8 percent capital adequacy ratio requirement, the equity holder can maintain up to $10 in debt, which gives a maximum activity level of $10.87.[120] The shareholder can, if she wishes, engage in more financial activity; all that is required is that she commit more of her own money. This is true whether or not the activity in question is socially benefical or costly. The question that must be asked of prudential regulation is: will she commit such another dollar, even if additional investment is unproductive and only increases risk?

The answer is, unfortunately, yes, given that the increase in risk is large enough. If there is an additional high risk project available (recall that these cost $1 and return $3 or -$1 with even odds), the shareholder's net increase in expected payoffs from adding this project is $0.50.[121] This is so even though the project is not, in expectation, profitable, and even though it increases expected social costs by $0.16. In contrast, if all that were

---

118

[119] Actual capital adequacy requirements are more complicated, as they require risk-weighting of the firm's assets, though the analysis remains the same.

[120] $.87 divided by 8% equals approximately $10.87.

[121] Recall that in the event of success, the shareholder will keep the gains in excess of the debt outstanding, while in a failure the shareholder simply walks away with $0.

available were an additional unprofitable low risk project, then the shareholder would be indifferent, since her net return from the investment is zero.

It is also worth noting which projects will be abandoned as capital adequacy ratios tighten. Suppose the regulator tightens the capital adequacy ratio from 8% to 30%. Given the shareholder's endowment of $2, the maximum possible size of the bank is $6.67 – so which projects will the shareholder choose to abandon? The hierarchy of projects abandoned would be low risk unprofitable projects, low risk profitable projects, high risk unprofitable projects, and high risk profitable projects, in that order. Hence, a tightening of capital requirements may lead to a greater equity investment and more high risk investment – or it could lead to an abandonment of low risk projects and a flight into high risk ones. One thing it does lead to is less banking. Consider the extreme case: even in an all equity bank (a capital adequacy ratio requirement of 100%), the shareholder may still take excessive risks from a societal perspective. What changes is that the size of the bank will be smaller: the bank funded only by the shareholder's $2 endowment will be able to make only $2 of loans (in high risk projects, of course). The overall risk level may be lower even though average project risk is higher; what this means is that the level of lending has contracted enormously, as has the money supply. Neither of these outcomes is good for the economy.

Risk weighting may alleviate some of these problems. Currently, however, the gradation of risk weighting is sufficiently coarse that it does not account for such differences among assets. For example, a bank issuing a high yield and high risk commercial loan would receive the same risk weighting as a low yield, low risk commercial loan of the same term. Nor does Basel III make distinctions among quality of home loans or among trading positions of similar maturities.[122]

This is not to say that capital adequacy rules are without value. What such requirements do is to keep the shareholders from expropriating the firm's debt holders. Every dollar that the shareholder puts in, keeping the assets constant, is a dollar that cushions the depositors and other creditors in the event of a loss.

---

[122] For a summary of the Basel III rules, see David Polk, "U.S. Basel III Final Rule: Visual Memorandum," July 8, 2103, *available at* www.davispolk.com%2Fsites%2Fdefault%2Ffiles%2Ffiles%2FPublication%2F3e28c060-fd34-42c0-9b75-003fe1c4ea5c%2FPreview%2FPublicationAttachment%2F937e4d31-4e86-4a00-a48b-00629c05ca4f%2FU.S.Basel.III.Final.Rule.Visual.Memo.pdf&ei=7H_6UtUr8sLIAbuXgcgI&usg=AFQjCNGvzL45Rfqy6do_k3PrqhkpcRlHzw&sig2=m4Lkli4h2QWiu5fXqDkDmw&bvm=bv.61190604,d.aWc&cad=rja.

Minimum capital rules may help mitigate losses otherwise borne by creditors and the FDIC, but it will do little to prevent excessive bank activities that generate systemic risk or other social harms. Such a rule makes sense given FDIC insurance (which renders depositors indifferent to banking risk), but it is a costly form of risk limitation.

iii. Financial transaction taxes

A tax is often preferable to a direct limit on activity. Pigovian taxes, in which tax rates for goods and services are tied to their expected social cost, can be a useful mechanism for forcing actors to internalize the costs of their behavior. For example, automobile use leads to externalities in the form of pollution and congestion. While one approach would be to limit the amount of driving that people may do, the hours at which driving may occur, or the number and type of cars that can be sold, such a mandate would fail to take into account private information about costs and benefits and, even with perfect information, would fail to satisfy heterogeneous tastes and needs.

A better approach may be to impose a tax upon automobile usage or gasoline consumption, with the tax level set at the level of externality. If the tax is set at exactly the externality level, then even individual users will make socially optimal choices.[123] While there may be some conceivable heterogeneity in externalized costs among end users, such differences are likely quite small relative to the common and uniform costs imposed (such as C02 emissions) by burning a gallon of gasoline. Such an approach is likely superior to the command-and-control approach or the no-reguation case.

So, then, by analogy, if financial transactions lead to systemic risk in the same way that gasoline consumption leads to pollution, a tax on financial transactions might make sense. In 2009, the Speaker of the House proposed such a tax, and France recently imposed one to apply not only in France but on all transactions involving French firms, including ADRs of French companies traded in the United States.[124] The idea, which

---

[123] This approach, known as Pigouvian taxation, has broad appeal among economists for precisely these reasons. Economist Greg Mankiw started the "Pigou Club" to collect the names of economists who support a broader use of Pigovian taxes. *See, e.g.*, Greg Mankiw, "Rogoff joins the Pigou Club," Greg Mankiw's Blog, Sept. 6, 2006, available at http://gregmankiw.blogspot.com/2006/09/rogoff-joins-pigou-club.html.

[124] *See* Thomas Ferraro and Andy Sullivan, *Wall Street Tax "Has a Great Deal of Merit": Pelosi*, REUTERS, December 3, 2009, *available at*

originated with John Maynard Keynes, is to reduce risk taking by taxing marginal trades on the belief that these are more likely to be based on "speculation" instead of "investment."[125]

But, there are some clear problems with such an approach, as the foregoing analysis of activity level limits suggests. As with other forms of ex ante regulation, the regulator must have a great deal of information available to it about the projects profitably available to each bank in order to set a reasonable tax rate. Setting the tax level too high prevents socially beneficial transactions, while setting the tax too low allows too much risky banking activity. There may also be a problem, again, with bank heterogeneity: if different banks have different sorts of opportunities with differnet risk profiles, then a one-size-fits-all approach will tend to punish banks whose activities carry little inherent risk and reward ones whose activities are of above-average riskiness. Proposed taxes attempt to differentiate among activities with, for example, different rates of tax for transactions in each of stocks, bonds, and derivatives.

This brings us to the fundamental problem of a transactions tax: the number of financial transactions or projects is a flawed proxy for systemic risk created by the entire bank portfolio. Such a tax may then be ineffective, or even have perverse effects, pushing banks into overall riskier transactions. The reason is that higher risk projects, which carry higher payoffs for the shareholder, are more likely to remain profitable under a tax regime.

Consider again our simple model. The bank is financed with equity of e = \$.80 and FDIC-insured deposits of \$0.20, such that the bank can engage in one project. The bank faces a choice between one high and one low risk project (recall that all projects are essentially coin flips, paying \$3/-\$1 and \$2/\$0 for high and low risk projects, respectively, plus profitability π). Suppose that the high risk project is not profitable ($n_h = 0$), and that the low risk project is profitable ($n_l = 1$), with expected profitability of $\pi_l = \$0.65$. Finally, the expected systemic harm from the high risk project is \$0.16, while that from the low risk project is \$0.08.

Without the tax and under a laissez faire system, the bank will invest in the profitable low risk project and eschew the unprofitable high risk one. Expected returns from underaking the

---

http://www.reuters.com/article/2009/12/03/us-usa-congress-tax-idUSTRE5B24J520091203 (last visited July 17, 2012).

[125] The idea for such a financial tax goes back to at least 1936, when John Maynard Keynes proposed it as a curb on excessive speculation. See Dr. Stephen Spratt, Intelligence Capital (September 2006). "A Sterling Solution". Stamp Out Poverty report. Stamp Out Poverty Campaign. pp. 15–16. (available at http://www.stampoutpoverty.org/?lid=9889; last viewed July 17, 2012).

low risk project are 50% * ($0.65 - $0.20) + 50% * ($2.65 - $0.20) - $0.80 = $0.65. Expected returns from undertaking the high risk project are 50% * $0 + 50% * ($3 - $0.20) - $0.80 = $0.60, which is lower. In this case, the laissez faire system achieves the social optimum, since the low risk project has private returns to the shareholder of $0.65, no externalization to counterparties or depositors, and systemic externalized costs of $0.08, for a net benefit of $0.57. Undertaking the high risk project, in contrast, would have private shareholder benefits of $0.60, expected losses to counterparties of $0.50, expected losses to depositors (the FDIC) of $0.10, and externalized systemic costs of $0.16, for a net societal loss of $0.16

Now, suppose the regulator attempts to be a good Pigovian by imposing a tax equal to the average systemic externality of projects, which in this case would be $0.12 per transaction. Perversely, this causes the bank to switch from the profitable low risk project to the unprofitable high risk project, making society worse off. Consider first the shareholder expected payoffs to the low risk project, which are lowered by the full value of the tax because of the bank's continued solvency given either success or failure: 50% * ($.65 - $0.12 - $0.20) + 50% * ($2.65 - $0.12 - $0.20) - $0.80 = $0.53. The high risk project, on the other hand, allows the shareholder to escape the incidence of tax in the failed state of the world: 50% * 0 + 50% * ($3 - $.12 - $.20) - $0.8 = $0.54. Payoffs to the high risk project are now higher for the shareholder (by a penny), causing the bank to switch under the tax regime to the high risk project. Social welfare is now $0.54 - $0.10 - $0.56 - $.16 + $0.12 = - $0.16.[126] Not only did the tax actually increase systemic risk, but it caused the loss of a profitable project. The total welfare loss under the tax regime is $0.70 relative to the laissez faire state.

A somewhat more general failing of a financial transactions tax is that the private payoffs are increasing in the degree of externality created. Unless the tax can be calibrated to the actual risk created, imposing a tax only deters small risks with small payoffs; larger risks are relatively undeterred. For example, suppose $\pi_l = \pi_h$ =$0.10 for all projects, while the systemic externalities of the low and high risk projects are given, respectively, by $\frac{\partial R}{\partial l}$ = $0.08 and $\frac{\partial R}{\partial h}$ = $0.16. This means that low risk projects are socially desirable, while high risk projects are not.

---

[126] The left hand side of the equation is, in order, the shareholder's expected private benefit, the depositor's (FDIC's) expected loss, expected externalities borne by counterparties, expected systemic harm, and the amount of tax collected by the government.

Now apply a Pigovian tax of $0.16 per transaction. Because the high risk project provides ample upside to the shareholder, high risk projects will be undeterred, even in an all equity bank.[127] Low risk projects will, however, not be undertaken, since the upside is simply too low to make up for the tax.[128]

For a financial transactions tax to work well, the regulator would have to link the tax directly to the expected social costs in each case, including costs put onto creditors and counterparties. These may all vary project by project, and bank by bank. This level of granularity is puts huge informational demands on the tax setter, and largely defeats the point of a Pigovian tax in the first place.

iv. Command and control – a financial FDA

As discussed above, part of the reason activity level limits and financial transactions taxation will not work is due to the heterogeneity of projects, and the fact that private rewards may often be correlated with externalized harms. Taking that into account, some proposals call for treating different sorts of financial transactions differently, perhaps allowing some and prohibiting others, based on their cost-benefit tradeoffs. This sort of command-and-control approach is analagous, for instance, to US regulation of energy production, where some forms of production are mandated or prohibited, or to regulation of new medical drugs, where a testing and approval process is necessary before new drugs may be marketed. This analogy is, in fact, explicitly used in a proposal by professors Glen Wyles and Eric Posner for a "financial FDA." They recommend financial products be approved before they can be sold, with the goal of eliminating or reducing needlessly speculatve (that is, risky) transactions that regulators believe will impose significant social costs.[129]

A potential stumbling block to this approach is that it is dubious that regulators can successfully identify, ex ante, which sorts of transactions impose unreasonable risks and little benefit. Finance is a fast evolving world where new products and markets arise constantly; keeping regulators up to speed on new developments would require an army of quantitative Ph.D.s in financial economics to throughly analyze every new financial

---

[127] The shareholder's before-tax expected payoff in an all equity bank undertaking the high risk project would be 50% * ($0) + 50% * ($3.10) - $1 = $0.55.

[128] The shareholder's before tax expected payoff in an all equity bank undertaking the low risk project would be 50% * ($0.10) + 50% * ($2.10) - $1 = $0.10.

[129] *See* Posner & Wyle, *supra* note 20, and accompanying text.

instrument. Indeed, given that the systemic risk depends on what other firms throughout the economy do, pre-approval even with the requisite manpower is a dubious proposition. As noted above, it is unlikely that any type of transaction is socially dangerous in the abstract, instead depending heavily on the quality of the bank engaging in the transaction, how many transactions the entire banking system is making (that is, the total activity level), and the correlation of risk across transactions. Unlike the regulation of drugs, where the risk to the individual consuming the drug is based solely on whether and how much of the drug the indiviudal takes, here the risks for the bank and the banking sector cannot be estimated in a vacuum or for an individual bank, but rather must be determined based on the choices and practices of hundreds if not thousands of other actors.

Even assuming that riskier projects may be identified ex ante, a financial FDA cannot be a complete solution. The reason is that the same level of portfolio risk may be constructed in different ways; if very risky instruments are outlawed, a firm may be able to create the same overall portfolio characteristics with an agglomeration of lower risk instruments.

Returning to our model, suppose that a bank has available to it both high and low risk investments, with a $1 equity investment by the shareholder. If the bank invests in a high risk project, its possible returns will be either $3 or -$1, each with probability 50%; expected payoffs to the shareholder are $0.50. Suppose now that high risk projects are prohibited under a financial FDA approach. Can the bank recreate the overall portfolio? Certainly: the bank can raise $1 in deposits, and invest in two low risk projects. This yields exactly the same risk-return profile as one high risk project: in the good state, the bank has gross returns of $4, of which it must pay $1 back to its depositors; while in the bad state, the bank has $0 and owes its depositors $1. The shareholder's expected return is the same, $0.50. Even though the financial FDA has prohibited the high risk project, little has changed.[130]

This is just a specific example of a more general point: modern financial instruments can create any risk-return combination in countless ways that defy easy categorization. A common illustration of this point is the concept of put-call

---

[130] Who bears the loss is different in the two cases. In the low risk case, the bank must raise deposits to purchase the two projects. With the high risk project, no deposits need be raised, and the subsequent losses are borne by the trading counterparty. Given the conventional-wisdom assumption that neither depositors nor trading counterparties adequately protect their positions from credit risk, this difference may not be material.

parity.[131] Using a mix of traditional and derivative products, any given asset return can be artifically created by combining different types of financial products, including buying assets, buying derivatives, borrowing, lending, and so on. In other words, whatever the financial FDA outlaws can likely be recreated through some combination of alternative financial instruments. The nimbleness of investors at avoiding regulation arises in other areas of law as well, where traditional categories of securities ownership have become increasingly problematic. [132]

We think it likely that banning particular instruments would do much good. The financial industry has been particuarly adept at regulatory arbitrage. As a poignant example, the multi-trillion dollar market for credit derivatives arose in part because regulations prevented certain entities, like insurance companies, from making certain types of investments.[133]

Put another way, the failing of a financial FDA is that it does not focus on overall portfolio risk. To do so, the financial FDA would need some complementary regulation, such as an activity level limit. For instance, if the regulator both proscribed high risk projects and limited the bank to just one project in total, the regulator would successfully limit aggregate risk. At the very least, the regulator's task will be more information-intensive than simply identifying a black list of prohibited instruments; rather, the

---

[131] *See* Alvin C. Warren, Jr., *Financial Contract Innovation and Income Tax Policy*, 107 HARV. L. REV. 460, 465–70 (1993); Fischer Black & Myron Scholes, *The Pricing of Options and Corporate Liabilities*, 81 J. POL. ECON. 637, 649–54 (1973); Robert C. Merton, *Theory of Rational Option Pricing*, 4 BELL J. ECON. & MGMT. SCI. 141, 141–42 (1973).

[132] Such issues arise elsewhere in the law, such as bankruptcy, tax, and corporate transactions. *See, e.g.*, Douglas G. Baird & Robert K. Rasmussen, *Anti-Bankruptcy*, 119 YALE L.J. 648 (2010) ("Investors today exploit the lessons of put-call parity, and they devise investment contracts narrowly tailored to their needs with various types of derivatives. Terms such as senior secured debt, junior secured debt, senior subordinated debt and junior subordinated debt are both common and carry little fixed meaning. Figuring out precisely what cash flow rights come with a certain investment often requires careful reading of a contract running for many pages. Far from having a relatively undefined right to the debtor's assets, these investors have particular and highly defined rights. The ownership interest is hardly as simple as holding a piece of senior or junior debt. Any particular investor holds a package of claims and derivatives.")

[133] State law generally prohibited insurance companies from investing in debt with a rating of less than a certain credit rating, say AAA. This cut insurance companies off from most of the corporate debt market, until financial innovators developing mechanisms for generating what they thought was AAA-rated debt out of the debt of many companies with less good credit ratings. *See, e.g.*, M. Todd Henderson, *Credit Derivatives are Not 'Insurance'*, 16 CONN. INSURANCE L.J. 1 (2009-2010); *see also, e.g.*, N.Y. Ins. Law § 1402(a) (2006); Nat'l Ass'n of Ins. Comm'rs, Model Law, Regulations, and Guidelines vol. III (2009) (describing risk-based capital levels).

regulator must come up with a way to prohibit certain types of overall portfolios.

What this analysis suggests is that it is not enough to identify highly risky activities: lower risk activities must be appropriately moderated as well, since they can be combined, either directly or synthetically, to create higher-risk activities. Thus, even if the regulator is able to observe, ex ante, high risk financial products, in order to impose meaningful limitations on social risk, the regulator must also be able to observe the productivity of each sort of transaction and set activity levels accordingly, across all actors in the financial system. It would be pointless, for instance, to disallow synthetic collateralized debt obligations (known as CDOs) and the credit default swaps that back them,[134] and yet allow accumulation of mortgages that may yield the same risky exposures. Given the innumerable permutations into which these sorts of financial instruments may be arranged, the financial FDA will have less information about these synthetic combinations and techniques ex ante than it will after observing them in practice. Overall, it may pay to wait.

4. A middle ground: the regulatory veto

The foregoing discussion of potential banking regulatory techniques reveals a general problem with all extant and proposed regulations: ex ante limits on the amount or sort of banking activity that may be undertaken are likely to be either ineffective at limiting risk or, if effective, costly because the regulator makes policy largely in the dark. Not regulating appears to be a poor option as well, since litigation and insurance will fail to constrain risk-taking activity due to problems of judgment-proofness. The problem, then, is that it is not tenable to wait until harm actually occurs, but neither is it workable to regulate ex ante what banks can and cannot do.

Fortunately, there is something of an intermediate solution in what we have termed the regulatory veto. Under the regulatory veto, the regulator does act to constrain banking activity, but operates after more information is gained by observing firm and customer choices, as well as getting better estimates of social losses. By refraining from ex ante regulation, the regulator avoids proscribing harmless or even beneficial activities; by retaining the right to observe the bank's activities and to shut down the bank

---

[134] This is, to an extent, what Congress has done in the Dodd Frank Act. *See* James Spindler, *Making the Next Financial Crisis Worse, One Regulation at a Time*, FORBES, Oct. 5, 2011 (describing the problems with the SEC release on synthetic securities).

before actual harm has been incurred, the regulator can maintain meaningful deterrence and avoid problems of judgment-proofness.

While this appears to be a great advantage of the regulatory veto, and we think explains its existence, any theory of the regulatory veto must consider why it is that banks were nevertheless encouraged to engage in excessive risk-taking activity. The regulatory veto existed before the Savings & Loan Crisis and before the Great Financial Crisis, and yet it was not effective at limiting excessive bank activities. Why did the regulatory veto, which we think is the best approach to banking regulation, fail?

As we develop in this section, the reason why is twofold. First, the structure of the regulatory veto encourages ultimatum-type behavior on the part of banks; for example, a bank may seek to become too big to shut down, or otherwise plan such that any shutdown will be excessively costly. Second, the incentives of regulator are wrong: being constrained to maximize ex post social welfare means that the regulator will find it difficult to make credible threats to shut down even a bank that provides no (or even negative) social value.

To develop the intuition, we start with a not-unrealistic hypothetical. Bank 1, operates under a regulatory veto regime in which the regulator acts to maximize social welfare. Knowing that the regulator acts only *after* the bank assembles a portfolio of loans and other investments, Bank 1 assembles an aggressive portfolio of loan assets and liabilities. Bank 1 then presents this portfolio to the regulator for inspection. While the portfolio is much more aggressive than the regulator would prefer, the loans have already made and the liabilities have already been incurred. The loans may even be profitable to the bank, which is powerful (albeit erroneous) evidence of their social value.[135] In effect, Bank 1 is daring the regulator to shut it down. The regulator's choice, therefore, is a starkly limited one: the regulator will shut down the bank (or the particular banking activity) if, and only if, society is made better off by shutting down the bank (or the activity). So long as Bank 1 presents a portfolio of investments with non-negative social value, it will be extremely difficult for regulators to shut it down.

A simple example shows why. If Bank 1's loan portfolio has an expected profit to Bank 1's stakeholders of $10, and an expected cost (borne by counterparties and society) of $9, the regulator maximizes social welfare by refraining from shutdown. Social welfare in that case is $1, whereas it would be zero in the

---

[135] Bank examination manuals explicitly instruct examiners not to be distracted by bank profits, but examiners of failed banks and thrifts commonly cited bank profitability as an excuse for regulatory forbearance. *See, e.g.*, Henderson & Tung, *supra* note 8 at 1024-25 (citing the example of WaMu).

event of shutdown. The optimal amount of bank activity might be strictly less than that which generates $10.00 for Bank 1, but obtaining this maximum may be difficult since the bank, as first mover and the bundler of investment projects, can control the portfolio of investments presented to the regulator.

In fact, Bank 1 could push things further, imposing costs on society of $9.99, such that social welfare is virtually a wash. Even in this case, the regulator will not veto the activity, since doing so would not increase social welfare. In fact, if the costs of shutdown are positive, then Bank 1 can present a portfolio of activities that have a social welfare of zero (or even negative, up to the costs of shutdown) and still have regulators forebear from shutting down the bank.

A simple picture of this dynamic is shown on Figure 3. As in our model, expected bank profits are increasing in bank activity levels, although the marginal benefit is decreasing beyond a certain point. (Importantly for setting the optimal regulatory policy, this point will be different for each bank.) Expected social costs of bank activities are also increasing, but unlike firm profits, the marginal social cost is constantly increasing in activity levels.[136] (Similarly, the social costs are likely different for each bank or each activity.) In this case, the optimal amount of bank activity (that is, the social welfare maximum) is where the sum of the two curves is at a maximum – this is shown as point A on Figure 8. This is the regulatory sweet spot, and the point where all regulatory reforms are trying to reach. Ex ante regulations are trying to estimate it; ex post attempts would be about discouraging banks from missing it. But, in light of the ultimatum game dynamic, a bank can increase its profitability (but not social welfare) by moving up the profit curve to point B, where its profits is at a maximum. It will do so, of course, only if it believes that the regulators will not shut down the activity. The regulator is in effect indifferent between points B and C on the curve, since in both cases the social welfare is zero. If shut down is costly (as it certainly is, along many dimensions), then regulators will not act to move to point C, despite the fact that social welfare could be increased (at the expense of bank profits) by moving down the profit curve to point A.

[136] If this were not the case, the social optimum would be one bank of incredibly large size.

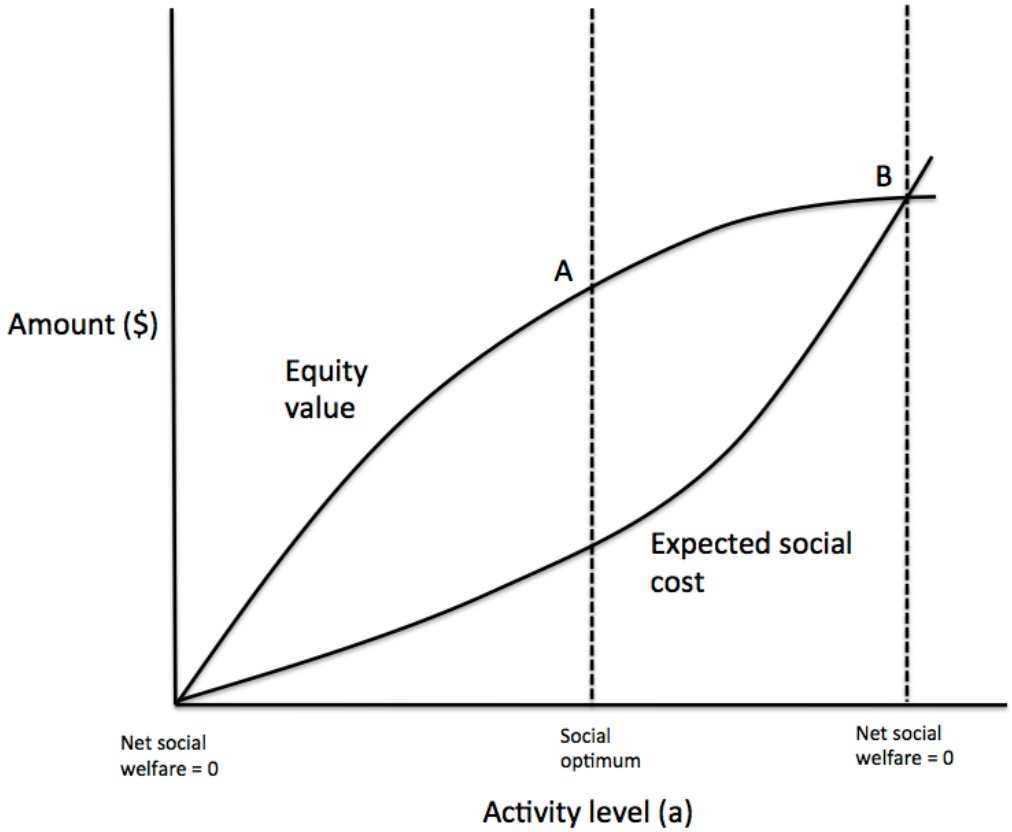

Figure 3: Model of bank profits and social costs.

This kind of behavioral dynamic was evident in the failure of Washington Mutual, which serviced nearly $125 billion in home loans and had over $200 billion in deposits prior to its collapse in 2008. Five years before it collapsed, examiners reported significant deficiencies in WaMu's underwriting process for residential loans, its core loan activity. But regulators brought no enforcement actions against WaMu, despite the fact that the problems persisted and even worsened. Examiners knew about WaMu's activities, they documented the excessive risk and lax controls, and were troubled by the amount of risk it was taking.[137] But they did nothing. When examiners were asked by an inspector general why they did not take action to shut down WaMu's excessively risky activities, examiners responded, "even though underwriting and risk management practices were less than satisfactory, WaMu was making money and loans were performing."[138] WaMu, in effect, presented its regulators with an ultimatum to condone their aggressive activity level, which was making large profits but at diminishing or negative amounts of social value, or shut down the

137
138

bank. WaMu chose point B, instead of point A, and effectively dared the regulators to shut them down.

The key weapon WaMu apparently deployed was the fact that its loans were profitable. Bank profits may be a particularly powerful weapon in an argument about shut down, since they may, to the uninformed, suggest social value. Regulators higher ups specifically informed bank examiners that "it is not sufficient to demonstrate that the loans are profitable or that the [bank] has not experienced significant losses in the near term."[139] This is because in theory bank regulators are aware that social welfare and bank profits are not the same, as seen clearly on Figure 8. But it may be much more difficult to politically shut down a profitable bank than an unprofitable one. This may be true as a matter of political rhetoric, as the history of banking regulation shows frequent and disastrous involvement by politicians in the regulation of specific banks.[140]

But it may also be true as a matter of existing law. Banking law statutes generally limit regulators ability to shutdown an entire bank to cases where the bank is "insolvent."[141] Although it is widely accepted that bank regulators have "wide latitude to define insolvency,"[142] this restriction may be unnecessarily hamstringing regulators into believing they have few options in the face of bank profits. This simply makes the dynamics of the ultimatum game worse than they would be if the regulators have more flexibility to shut down entire banks. We will return to this issue below in Part V, where we discuss potential reforms.

We return now to our simple model of bank behavior to illustrate explicitly these benefits and drawbacks of the regulatory veto. As before, the bank chooses how many projects it undertakes, starting with the profitable ones and moving on to those that do nothing but increase risk. Instead of imposing ex ante limits on the bank, the regulator waits for the bank to act; while waiting, the regulator observes a signal of the firm's expected production and the expected systemic cost $R(a)$, where, for concreteness, $a = 2l + h$. After observing these signals, the regulator can either allow the firm to continue, getting the expected

---

[139] *See* Dept. of Treasury, Offices of Inspector General, "Evaluation of Federal Regulatory Oversight of Washington Mutual Bank," Report No. EVAL-10-002, Apr. 2010 ("WaMu Report") at 20 (citing guidance from WaMu's bank regulator, the Office of Thrift Supervision).

[140] For a discussion of how political influence impacted regulatory action during the S&L Crisis, see Epstein & Henderson, *supra* note 37, at 516-517; 536-539.

[141] *See* FDIC Report, "An Examination of the Banking Crises of the 1980s and Early 1990s," ("FDIC Report") available at www.fdic.gov/bank/historical/history/421_476.pdf at 457.

[142] *Id.*

benefits of production and incurring expected systemic costs, or the regulator can shut down the firm, forcing a liquidation that yields $d + e - C$, where C is the cost of liquidation and, as before, $d + e$ is the capital investment of debt and equity holders; that is, a shutdown and liquidation is able to recover the initial capital investment, but at some significant cost.

Perfect Information Case

For simplicity, we start with the case where there are no agency costs, shareholders have perfect information, and the regulator's subsequent signal is perfectly revealing. Reasoning by backward induction, shareholders will choose the maximum level of activity that does not lead the regulator to shut down the firm or the activity. Shareholders know that the social welfare-maximizing regulator will shut down the firm or the activity if and only if the sum of the expected payoffs to equity, debt, and society, plus the cost of shutdown, exceeds zero. With this in mind, shareholders choose the level of activity that leads to the regulator being indifferent to the bank's continuing. Formally, shareholders choose h and l such that $\sum_{i=l,h} n_i \pi_i - R(a) = -C$.

The good part of this approach is that all profitable projects are undertaken and the firm limits its overall riskiness to an ex post acceptable level. The drawback of this approach is that, because of the ultimatum-game dynamic with the regulator, the overall benefit created by banking is actually negative: it is $-C$, to be specific. Society would be better off without it. This is so even when we assume that the regulator receives perfect information with regard to risk, profitability, and social cost. The problem becomes worse if we imagine that shareholders have the ability to affect the cost of liquidation, C. As C grows larger, so too does the level of risk that shareholders will choose, and social welfare decreases in a one to one correspondence.

There are many ways to increase C, and we should expect banks, all else being equal, to utilize them to increase the chances that it can maximize its own profits. For instance, banks will tend toward complex transactions that are difficult to understand and unwind. In addition, banks will increase interconnectivity with other banks, which makes shut down of particular activities more difficult and costly. They will also use separate legal entities, such as subsidiaries and special-purpose vehicles, especially ones outside of the jurisdictional reach (either physically or statutorily) of particular regulators, so that corporate fictions will make shut down messier. Capturing regulators, either explicitly or implicitly

through persuasion, is another mechanism for increasing the costs of shut down. It is more difficult to hurt individuals that one knows, likes, and works with on a regular basis. Deploying politicians to give regulators a hard time, such as with hearings on Capitol Hill, is another tried and true mechanism. Finally, we should expect banks to certainly take no ex ante steps to ease subsequent liquidation, as is supposed to be the case with so-called "living wills," since this simply lowers the costs of shut down, and therefore decreases expected bank profits in expectation.

Imperfect Information Case

In the case where the bank and regulators have imperfect information, the social welfare may actually be higher than when information is perfect. This is so because the bank is less sure of the regulator's signal and would not wish to risk complete shutdown and loss of all profits. In this case, the bank may error on the side of less profits, and, as it moves down the profit curve toward point A in Figure 8, increasing social welfare. More formaly, suppose that the regulator observes the amount $\sum_{i=l,h} n_i \pi_i - R(a)$ only with precision of plus or minus \$1, relative to the firm's signal. There will then be cases in which it behooves the firm to set a portfolio of activities such that $\sum_{i=l,h} n_i \pi_i - R(a) = \$1$, which represents an increase in social welfare of \$1 over the perfect information case. This is not to say, however, that less information is necessarily good; one could imagine the extreme case where the regulator's signal is so uninformative that it then behooves the bank's shareholders to maximize both activity level and risk (although, ex ante, they may be less willing to invest in banking firms).

* * *

At this point, we have largely set out to do the work we intended to do. We have shown, with some simple intuition and a more formal model that the extant and proposed mechanisms for optimizing bank risk taking are insufficient to accomplish this task. We have shown why regulators can do better by avoiding ex ante regulation, and waiting until they have both a signal as to the value of bank activities to the bank and its customers, and a signal about the social costs of bank activities. But we have also pointed out a fundamental flaw in the regulatory veto approach. Although

we leave it to others and other work to propose better fixes to the banking regulatory system, in the next Part we briefly offer some thought starters on potential ways to improve banking regulation.

IV. How to Make Banking Regulation Better

As should be clear at this point, we are not optimistic that any form of ex ante or ex post regulation alone is going to help achieve the optimal amount of bank activity levels. The current reforms miss out on an important insight: the benefits of information acquisition and processing are increasing in time, which suggests there is benefit of delaying regulation, while the costs arising from judgment-proofness are also increasing in time, which suggests there is benefit in accelerating regulation. Regulatory action is optimal where the tradeoff between these is at its maximum, suggesting the best approach is to improve the operation of the regulatory veto, rather than to focus on either ex ante or ex post controls. Improving the dynamic between bank and regulator as it involves the onsight examination and regulators ability to fine tune bank activities is where regulatory reform can have the best effect.

Fred Tung and one of the authors have offered two ideas for giving regulators better incentives to get tough with banks, to disregard bank profits in the face of socially inefficient risk, and to try to march banks down their profit curve toward point A in Figure 8. First, they proposed paying examiners for "performance," defined as optimizing bank regulation toward point A.[143] Although this proposal suffers from the problems around designing an optimal compensation package for bank executives discussed above, the idea of trying to develop a compensation scheme that encourages regulators to act where they currently forebear is consistent with what we've said here. One can simply think of incentive pay for examiners as a mechanism for overcoming the costs of shutdown (noted as "C" in the discussion above).

Second, they proposed an auction mechanism for allocating examiners to particular banks as a way of ensuring better fit, of encouraging greater production of information about bank quality and risk, and as a way of combating regulatory capture.[144] Again, this proposal is consistent with our argument, in that it is about reducing the costs of regulatory action, encouraging better accountability on the part of regulators, and improving examiner quality and work.

---

[143] See Henderson & Tung, *supra* note 8, at __.
[144] *See id.*

There are other possibilities, which we will only mention here. In general, our model suggests the best reforms are the ones that focus on reducing the costs of regulatory action (that is, reducing "C"), as the Henderson & Tung proposals try to do.

One of these might be the requirement in the Dodd-Frank Act that the largest 100 financial institutions (that is, those with greater than $250 billion in assets) submit "living wills" to bank regulators to assit the regulators in winding down the banks in the event of a failure.[145] In theory, such living wills could help lower the regulatory shut down cost for an activity or an entire bank. This could lead to more regulatory action in the form of the regulatory veto, and this might move us closer to the regulatory sweet spot. In addition, as noted above,[146] banks do not have incentives to make the costs of their shut down lower, and therefore a regulatory requirement to do so may make sense.

There are problems, as the rule now stands. First, much of the information contained in the first batch of living wills was repetitive of information about banks already publicly available.[147] Bank regulators already know this information, plus much more, so packaging only publicly available information in a new form is unlikely to add much value. Second, as currently formulated, such living wills are designed to be deployed only in extreme circumstances, such as the rapid and unexpected failure of an entire bank.[148] This is much less useful in terms of our model than a more general playbook on how to shut down specific bank activities, both as a stand alone matter and in combination. For obvious reasons, banks will not voluntarily provide this information, especially if it is made public. Our analysis suggests that it would be valuable for regulators to have much greater transparency and a road map to reduce the costs of shutting down bank activities (in our model, lowering C). This is not the current approach.

Another possibility is the use of a device through which regulators precommit to take a regulatory action in the event of

---

[145] Section 165(d) of the Dodd-Frank Act of 2010 requires certain large financial institutions to prepare and update plans for orderly resolution of their balance sheets in the event of a threatened or actual collapse of the firm. The rules are promulgated as Regulation QQ for the Federal Reserve Board (*12 C.F.R. pt. 243*) and Part 381 for the FDIC (*12 C.F.R. pt. 381*). For a list of and access to resolution plans required under the law, see http://www.federalreserve.gov/bankinforeg/resolution-plans.htm.

[146]

[147] *See* Jessica Silver-Greenberg & Nelson D. Schwartz, *"Living Wills" for Too-Big-to-Fail Banks Are Released,* N.Y. TIMES, July 3, 2012, *available at* http://www.nytimes.com/2012/07/04/business/living-wills-of-how-to-unwind-big-banks-are-released.html.

[148] *Id.*

certain eventualities. In other words, regulators would remove some of the discretion from the examination process, so as to signal to banks that the costs of regulatory action are low, and therefore try to alter the incentives in the regulatory ultimatium game that is currently played. If banks believe regulators will exercise the regulatory veto in certain circumstances, then when they move first, they will be be reluctant to generate those circumstances. Of course, predicting the optimal circumstances in which to exercise the veto is a difficult challenge, since this falls int the ex ante bucket of regulation, and therefore is subject to the criticisms we've outlined above. But the virtue of combining an ex ante precommitment device with the regulatory veto, is that it allows the regulators to observe costs and benefits, and only shut down those activities that, *in practice*, result in suboptimal levels of bank activity or care.

An analogy to this regulatory approach are the so-called "prompt corrective action" rules (PCA).[149] Put in place after the evidence that politicians influenced regulators to forebear from taking action during the Savings & Loan Crisis, these rules remove regulator discretion to act when certain triggers are met.[150] For instance, if bank leverage ratios fall below prescribed levels, regulators have no choice but to act to, among other things, force the bank to raise additional capital. This approach could be expanded to include not only compliance with leverage limits, but also to encourage optimal bank activity levels. Regulatory higher-ups could estimate the optimal number of regulatory curtailments based on historical practice, prevailing bank practices, and macro-economic conditions. A target could be set, with an algorithim for determining whether regulatory action is warranted in a particular case. The key insight of the PCA model is that regulators may be unable to act when it is optimal for them to act. In the PCA case, the disease was meddling by politicians, who were beholden to bankers. In this case, it is regulators subjected to an ultimatum game in which bankers can force regulators into a rational but suboptimal regulatory choice.

V. CONCLUSION

Banking regulation has repeatedly failed, most recently in the run-up to the Financial Crisis. The solutions offered for the latest failures largely fall into traditional categories: ex post remedies (following from tort law) and ex ante prophylaxes

---

[149] Section 38 of the Federal Deposit Insurance Act, 12 USC § 1831o.
[150] *See* Epstein & Henderson, *supra* note 37, at 516-17.

(following from command-and-control regulations, like in environmental law). Proposals to change banker incentives through mandatory forms of compensation, to limit the size of banks, and to require additional capital all fall into these categories. While individual reforms might be somewhat helpful, especially in combination, our analysis shows the fundamental weakness of relying solely on ex ante or ex post forms of regulation.

Ex ante regulations put burdens on regulators that are too high in terms of information and knowledge relative to the case where the same regulatory decisions would be made later in time. Regulators deciding on the optimal amount of bank activity or bank care simply do not have all the information they will need to make the best regulatory decision they can make. It might be better to wait. If regulators can learn about the private and social benefits and costs of particular bank activities by watching them in practice, then there are gains from regulatory delay. As we show, this is especially true in modern finance where there are heterogeneous firms, products, and consumers. In this case, an ex ante, one-size-fits-all approach will over- and under-deter conduct. In short, ex ante regulation might be sensible where expected harms are large, innovation is uncommon, and regulatory arbitrage is costly, but banking is (increasingly) an area in which while expected harms my be large, there is a tremendous amount of innovation and the ability to arbitrate regulations is extremely easy.

Delaying too much, however, may result in a lost opportunity for any effective regulation. While ex post litigation (or insurance) might be effective in producing optimal incentives for most consumer products and services, the nature of systemic risk means that for banks there are significant judgment-proofness problems, especially if defaults are highly correlated across the financial sector. If excessive bank activity is only revealed when banks fail, then tort liability does little good, especially with the presence of limited liability and (implicit and explicit) government guarantees. In such cases, the tort literature suggests insurance (or bonding) is efficient, but if the default risk of insurance companies is correlated with bank systemic risk, as appears to be somewhat true from recent experience, then ex post insurance will be insufficient to optimize bank risk taking.

It is for these reasons that we think bank regulation has historically involved another, more unique form of regulation, which we call the “regulatory veto.” Regulators who work at specific banks full time observe the banks’ activities, and are empowered to require a change in the activity level or care a particular bank takes. The logic is to try to optimize the tradeoff between information production and analysis by the regulators, which is

increasing over time, and the efficacy of any regulatory action, which is decreasing over time. The regulatory veto is designed to find this point, which is likely to vary by bank, by financial product, and depending heavily on the prevailing economic circumstances.

Unfortunately, we show how the current design of the regulatory veto is unlikely to result in its effective use. Because banks act first, they can present a socially inefficient portfolio of bank activities to regulators, who cannot improve social welfare by shutting down the activities. In other words, the current regulatory veto is inherently flawed by an ultimatum-game dynamic in which regulators rationally act in ways that destroy social welfare. It follows that banks will engage in practices, like opacity, complexity, and regulatory capture that are designed to influence this process.

We do not offer a complete solution for this problem, leaving that to others or other work. But our analysis does suggest that the reforms most likely to be effective at optimizing bank activity levels are ones designed to increase the ability of regulators to pre-commit to take regulatory action in the face of evidence of excessive bank risk taking.