Maryland Law Review

Volume 72 | Issue 4 Article 13

A New Perspective on the Costs and Benefits of Financial Regulation: Inefficiency of Capital Intermediation in a Deregulated System

Wallace C. Turbeville

Follow this and additional works at: http://digitalcommons.law.umaryland.edu/mlr

Part of the Banking and Finance Commons

Recommended Citation

Wallace C. Turbeville, *A New Perspective on the Costs and Benefits of Financial Regulation: Inefficiency of Capital Intermediation in a Deregulated System* , 72 Md. L. Rev. 1173 (2013)
Available at: http://digitalcommons.law.umaryland.edu/mlr/vol72/iss4/13

This Article is brought to you for free and open access by the Academic Journals at DigitalCommons@UM Carey Law. It has been accepted for inclusion in Maryland Law Review by an authorized administrator of DigitalCommons@UM Carey Law. For more information, please contact smccarty@law.umaryland.edu.

A NEW PERSPECTIVE ON THE COSTS AND BENEFITS OF FINANCIAL REGULATION: INEFFICIENCY OF CAPITAL INTERMEDIATION IN A DEREGULATED SYSTEM

WALLACE C. TURBEVILLE*

ABSTRACT

The evaluation of the costs and benefits of limiting human activity through regulation is at the core of our government. When the activity that would be restricted is commercial, many have a bias against regulation. They assume that restricting activity reduces productivity in the economy to the detriment of wealth creation. When a politician says that a rule is "unduly burdensome," he or she means that the discernible benefits do not outweigh the wealth that would be created if the rule did not exist. Today, the most sophisticated and powerful articulation of this point of view concerns the regulation of the financial markets through financial markets reform under existing and potential future legislation.

This Article describes how this debate suffers from preconceptions that rely on unfounded economic theory and misconceptions based on the sheer complexity of modern finance.

It identifies the first step: defining how to measure the value to society of a market structure that is bounded by a given set of regulations. The primary function of financial markets in this context is the intermediation of capital investment. Efficiency can be measured by cost. If the economic rent extracted by the financial sector for facilitating capital intermediation in a market construct is proportionate to the value added by this activity, efficiency of the market structure is demonstrated.

This Article postulates that this has not been the case in recent years. In doing so, it rejects the predominant approach to measuring efficiency that emphasizes the cost of individual financial transac-

Copyright © 2013 by Wallace C. Turbeville.

* Wallace C. Turbeville is a lawyer and financial services expert who has worked in the financial services industry since 1978. He practiced law for seven years and then joined Goldman Sachs as an investment banker specializing in infrastructure finance and public/private partnerships. He is now a Senior Fellow at Demos, a nonprofit advocating reform in the financial services industry. DEMOS, http://www.demos.org/.

tions. The predominant approach implicitly assumes that capital intermediation can be viewed as a linear path of a unit of investment from an investor to a productive use. In this view, the costs of individual transactions along the way determine the cost of the capital intermediation process.

This Article makes the case that this is naïve or intentionally misleading. Modern capital and derivatives markets are exceedingly complex and involve multiple methods for extraction of value by the financial sector that must be paid for by the productive economy. This Article postulates that the amount extracted is demonstrably far higher than historic data and reasoned analysis suggest could possibly be reasonable. Therefore, the rents extracted by the financial sector for intermediating capital investment are inefficiently high.

Finally, this Article asserts that the costs and benefits of financial market regulation should not be biased by the assumption that the restriction of activities is a burden that must be offset by legitimate benefits such as safety and soundness of the system or even price transparency. Under the proposed approach, regulations that reduce complexity or market power do not by definition burden the financial system in providing its essential social function but very likely enhance the efficiency of that process.[1]

INTRODUCTION

The debate that rages over regulations in the United States represents a long struggle to define the optimal role of government in society. The debaters question what behaviors should be permitted or forbidden and what yardstick should be used to answer these questions. Regulation of commercial activity impacts the economic performance of the country and directly affects prosperity.

Financial reform responsive to the financial crisis of 2008 and its role in the Great Recession is an important venue for that debate. The Dodd-Frank Act of 2010 was historic in its breadth and the scale of its changes.[2] Implementation of its provisions requires the promulgation of 236 separate rules, fewer than half of which have been completed as of this writing, more than two and a half years after en-

1. This Article is adapted from Wallace C. Turbeville, *Cracks in the Pipeline: Restoring Efficiency to Wall Street and Value to Main Street,* DEMOS (Dec. 5, 2012), http://www.demos.org/publication/cracks-pipeline-restoring-efficiency-wall-street-and-value-main-street.

2. Dodd-Frank Wall Street Reform & Consumer Protection Act, Pub. L. No. 111-203, 124 Stat. 1376 (2010) (codified in various sections of 12 U.S.C.).

actment.[3] Dodd-Frank is very different from the New Deal's financial regulation, largely because of the complexity of today's financial markets. It identifies some activities that are prohibited as unacceptably risky. For example, insured banks are prohibited from engaging in many types of proprietary trading activities under the "Volcker Rule."[4] But, overwhelmingly, Dodd-Frank improves existing markets rather than changes them fundamentally, focusing on transparency (limiting the over-the-counter trading of derivatives); the upgrade of risk management procedures (required clearing of many derivatives, capital and margin requirements, and post-trade data reporting for financial firms); and oversight by prudential regulators (standards for the amounts and adequacy of bank capital and required living wills to resolve failed institutions).

Regulation of the financial markets establishes boundaries within which these markets allocate capital to productive uses. In this way, Dodd-Frank's influence on the shape of the economy is strong. In crafting the law, Congress implicitly weighed the costs and benefits of adopting the market structure that results from Dodd-Frank against alternatives, including the alternative of doing nothing.[5] The agencies responsible for crafting implementing regulations are legally required to take this weighing of costs and benefits to the next level of detail, often by explicit statutory provisions.

As the agencies complete their work on the implementing rules, the next phase of the battle between the Goliath of the bank lobbyists and the David of the reform community will take place in the courts, particularly in the U.S. Court of Appeals for the District of Columbia Circuit. The first skirmish involved the Securities and Exchange Commission's Proxy Access rules.[6] The D.C. Circuit found the SEC's "cost/benefit" analysis to be inadequate.[7] This murky issue of costs and benefits promises to be the target of judicial review of all important financial reform rules.

While the courts consider the financial reform regulations, Congress continues the weighing process. Those expressing concern over "burdensome" regulation are proposing to limit the Dodd-Frank legis-

---

3. U.S. GOV'T ACCOUNTABILITY OFFICE, GAO 13-195, REGULATORS HAVE FACED CHALLENGES FINALIZING KEY REFORMS AND UNADDRESSED AREAS POSE POTENTIAL RISKS (2013).

4. Dodd-Frank Wall Street Reform & Consumer Protection Act, Pub. L. No. 111-203, 124 Stat. 1376, § 619 (2010) (Volcker Rule to be codified at 12 U.S.C. § 1851(h)(4)).

5. *Id.*

6. Bus. Roundtable & Chamber of Commerce v. Sec. & Exch. Comm'n, No. 10-1305 (D.C. Cir. filed July 22, 2011).

7. *Id.*

lation,[8] while others advocate broadening the scope of the law.[9] In both cases, the proponents are dissatisfied with the balance of costs and benefits in the structure of the financial markets.

From its inception, this debate over the best financial market structure in the post-financial-crisis world has avoided some basic questions. Ground rules for measuring costs and benefits have been elusive and methodologies have been inconsistent. As a result, a troubling assumption has lurked in the background. There is general consensus that a recurrence of the consequences of the near collapse of the financial system in 2008 must be prevented. But it is also fundamentally assumed that constraining the financial markets will impede economic productivity. Financial market regulation, therefore, is considered to be an inherently costly undertaking where the benefits (such as reducing the risk of financial crises) perhaps make it worthwhile.

The regime established by Dodd-Frank understandably gave priority to risk mitigation in response to the actual events of 2008. Massive risk is an element of the modern financial system. But risk and reward are generally symmetrical. An economy disproportionately skewed toward rewarding the financial sector raises important structural issues that have yet to be addressed by laws and regulations. The first step is to erase the preconception that regulations burden the efficient functioning of a market and that the good that regulations do must outweigh this burden.

This Article will demonstrate that the concept of definitionally efficient unregulated trading markets is fundamentally flawed. At its core, it is based on an incorrect measure of efficiency which leads analysts to look in the wrong places when measuring frictions embedded in market structures and behaviors. In the current discourse, efficiency is almost uniformly measured by referencing the cost of individual transactions. But the principal social value of financial markets is not to assure the lowest transaction costs for market participants. Rather, it is to facilitate the efficient deployment of funds held by investors to productive uses. In other words, markets are efficient if the cost to the entity putting capital to work productively is as close as possible to the price demanded by the entity that seeks a return on its investment, both measured comprehensively. The entire difference

8. *See, e.g.*, H.R. 1062, 112th Cong. (2012) (proposing the "Burdensome Data Collection Relief Act").

9. Letter from Rep. Keith Ellison to Ben Bernanke and Timothy Geithner (Sept. 24, 2009), *available at* http://ellison.house.gov/images/stories/Documents/2009/Letter_to_Bernanke_and_Geithner.pdf.

between the two is attributable to the plumbing that connects capital sources to capital uses, known as intermediation. The economic rents[10] extracted by intermediaries must be as low as possible to compensate them for performing the essential intermediation service if the system is to work efficiently.

Almost universally, this concept is lost in the discussion of financial markets. Efficiency is expressed in terms of the cost of a securities or derivatives transaction. This measures how well the markets work for traders. But it is only one element of the cost of intermediation between capital sources and uses. For reasons ranging from ideology to analytic sloth, the possibility that a market with low transaction costs can also be one in which intermediation costs are inefficiently high is ignored in public debate and academic analysis. As the courts begin to review cost/benefit analysis for the massive number of rules promulgated by the regulatory agencies, it is important that they be directed to properly measured costs and benefits.

Properly measured, the financial markets have become less efficient in the era of deregulation even though conventional wisdom dictates that advances in information technology and quantitative analysis should have caused the opposite result. Enormous sums of money are extracted from the capital intermediation process causing the financial sector share of the economy to grow at the expense of the productive manufacturing and service sectors and public finance. This trend must be reversed if the U.S. economy is to prosper and compete successfully in the world markets.

Several factors contribute to this result. Contrary to commonly held beliefs, advances in information technology and quantitative analysis have actually created asymmetries in information among trading market participants. While up-to-date information related to fundamental value (for example, corporate financial reports, crop yields, government policies)[11] is widely known today, these advances have been used by the more sophisticated and better-funded market participants to detect, analyze, and often influence activities by other mar-

10. A measure of market power, economic rent is the value in excess of marginal costs extracted by market participants. *See Economic A to Z Terms*, ECONOMIST, *available at* http://www.economist.com/economics-a-to-z/r#node-21529784 (last visited June 21, 2013).

11. Fundamental value refers to the intrinsic value of a stock, bond, or derivative based on available information. In the context of a share of stock, "Fundamental analysis entails the use of information in current and past financial statements, in conjunction with industry and macroeconomic data to arrive at a firm's intrinsic value." S.P. Kothari, *Capital Markets Research in Accounting* (Mar. 2001) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=235798.

ket participants, and to then exploit advantages derived from this market non-fundamental information.[12] In addition, complex instruments—primarily derivatives—are better understood by the financial institutions that market them than by their customers. As a result, the financial institutions profit far more from the sale of these instruments than their customers realize.

These two types of information advantage (and others) are systematically used to increase economic rents extracted from the intermediation process. Changes in both the law governing banks and the underlying structure of the financial markets have allowed banks to exploit this informational advantage through capital intermediation. The abandonment of the Glass-Steagall Act,[13] which gave rise to multifunctional universal banks, as well as the elimination of several large banks during the financial crisis, led to dramatic consolidation in the financial service industry.[14] The financial sector is now dominated by a small number of large banks that enjoy tremendous market power.[15] Because of powerful shared interests in the structure and process of the markets, these banks act as an oligopoly.[16] Concentrated market power allows the oligopoly to use its information advantages and massive capital to extract value from the intermediation process on a large scale. Dominant financial institutions systematically create market distortions and then exploit those distortions. Changes in the financial markets have similarly increased opportunities for capital intermediation. The growth of pooled investment vehicles, from pension funds to hedge funds to money market mutual funds and others, has changed the process of capital intermediation.[17] Much of the money that historically funded bank deposits has migrated to those vehicles. The bank lending model for intermediation—

12. The information asymmetry discussed in this Article is very different from the fundamental information inefficiencies addressed in a number of studies that examine the differing motivations of market participants. For an excellent analysis of such fundamental information inefficiencies, see David A. Hirshleifer, *Investor Psychology and Asset Pricing*, 56 J. FIN. 1533 (2001); Lynn Stout, *Inefficient Markets and the New Finance*, J. FIN. TRANS. (forthcoming), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=729224.

13. Glass-Steagall Act, Pub. L. No. 73-66, 48 Stat. 162 (1933).

14. Sergei M. Guriev & Dmitriy Kvasov, *Imperfect Competition in Financial Markets and Capital Structure* (Aug. 2007) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=603721.

15. *See* COMPTROLLER OF CURRENCY, U.S. DEP'T OF TREASURY, OCC'S QUARTERLY REPORT ON BANK TRADING AND DERIVATIVES ACTIVITIES FOURTH QUARTER (2010).

16. For an analysis of the post-Glass-Steagall oligopolistic behavior of banks in the debt and equity markets, see Guriev & Kvasov, *supra* note 14.

17. *See* WILLIAM H. DONALDSON, SEC. & EXCH. COMM'N, IMPLICATIONS OF THE GROWTH OF HEDGE FUNDS (2003) (discussing the rise of pooled investment vehicles like hedge funds).

where depository banks move funds held as customer savings into capital investments—has been largely replaced by a trading market intermediation model. In most cases, managers of pooled funds are judged by comparing transaction results with overall short-term market results rather than by long-term results.[18] The driving goal is to beat the market rather than to produce long-term results. Since long-term results are not emphasized, efficient intermediation between capital sources and uses is less valued by investors.[19] This obscures inefficiencies from both the investors and consumers of capital who could discipline the system to increase efficiency if performance were measured by giving greater consideration to long-term growth of value. It enables the extraction of value from the capital intermediation process by financial institutions.

The consequences of extraction of value from capital intermediation by the dominant financial institutions reach far beyond unfairness among market participants. Inefficient capital intermediation exacts wide-ranging costs that severely burden the economy by restricting the flow of capital to consumers, increasing the cost of consumer goods, increasing the costs of productive projects, slowing job creation, and reducing the effectiveness of monetary policy. Inefficient capital intermediation also reduces the self-correcting effect of supply and demand and excessively diverts wealth to the financial sector. The net effects of these costs are widely apparent: lower productivity, more expensive goods, and fewer jobs.

In recent years, the financial sector share of aggregate gross domestic product ("GDP") has been in the range of 8.3%, an increase from the historic level of 4.1%.[20] By inferring that the historical increase in financial sector share of GDP is attributable to the value diverted from capital intermediation, the excessive wealth transfer to the financial sector is in the range of $635 billion per year.

These market inefficiencies also increase the likelihood that financial crises, like the 2008 financial crisis, will recur. Exploitation of market distortions is the root cause of the most recent financial crisis. Market participants with enormous market power are incentivized to increase the complexity of markets and to take excessive risks to reap

18. Robin M. Greenwood & David S. Scharfstein, *The Growth of Modern Finance* (July 1, 2012) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2162179.

19. *Id.*

20. Justin Lahart, *Number of the Week: Finance's Share of Economy Continues to Grow*, WALL ST. J. (Dec. 10, 2011), *available at* http://blogs.wsj.com/economics/2011/12/10/number-of-the-week-finances-share-of-economy-continues-to-grow/.

short-term gain, often by generating asset or debt bubbles. Financial crises can be triggered by bursting price bubbles.

If these root causes are not addressed by regulatory reforms, future financial crises are inevitable. After all, the banks themselves could have restrained their behavior prior to the crisis even if regulations did not. It should be obvious to all that the incentives to the banks in terms of outsized corporate profits must be immense to induce behavior that put the bank at mortal risk. The lure of immediate profits from the exercise of market power was simply too strong to allow rational prudence to prevail.

I. UNDERPINNING OF THE GREAT DEREGULATION EXPERIMENT

Commencing with President Reagan's election in 1980, the financial markets in the United States were totally transformed by three major developments: (1) advances in quantitative measurement of dynamic market price movements and information technology; (2) concentration of market power into a few large financial institutions; and (3) deregulation.

The first development was the ability to quantify the future risks embedded in the market, apart from changes in fundamental value.[21] Instead of looking at economic data and financial statements, which are components in today's price, traders could determine the statistical probability of price moves going forward. They could ascribe a value to future price volatility and detach it from the underlying security or commodity. This was a precondition to the changes of the last three decades, and was essential to the rapid emergence of the $60 trillion per year derivatives market, a market that barely existed before the 1990s.[22]

Advances in information technology allowed near-real time valuation of dynamic price movements and deployment of capital at high speeds to exploit this capability. Previously, trading took place most often by telephone and was driven primarily by broad-based information relevant to macro-economic data and information relevant to the fundamental value of specific securities and commodities. Trading floors are now dominated by arrays of screens at every trader's

21. For a definition of fundamental value, see *supra* note 10.

22. Derivatives are contracts that obligate payment of cash based on asset market price movements. The advent of derivatives allowed traders to detach price movement risk from assets and the Black-Scholes model allowed the resulting instruments to be valued. *See* Wallace C. Turbeville, *Derivatives: Innovation in the Era of Financial Regulation*, DEMOS (June 13, 2013), http://www.demos.org/publication/derivatives-innovation-era-financial-regulation (discussing the history of and theory behind the Black-Scholes model).

desk, most tied directly to trading venues enabling instantaneous execution of trades. Trading was transformed into an electronic business of massive volumes and complex strategies intended to profit from dynamic price movement rather than simple investment in assets based on their fundamental value. Furthermore, the critically important regulatory framework that had defined the markets since the New Deal was utterly dismantled to remove all impediments to the new trading business.

Commercial banking (taking deposits and making loans) had been separated from investment banking (proprietary trading of securities and commodities) since the adoption of the Glass-Steagall Act in 1933.[23] Banks that took in deposits and made loans were prohibited from trading securities and commodities.[24] Throughout the 1980s and 1990s, regulators repeatedly undermined the separation of commercial and investment banking. The law was finally repealed in 1999 by the Gramm-Leach-Bliley Act.[25] All banks could now trade for their own account and the race was on to accumulate vast stockpiles of financial assets to increase the market power generated by marrying-up trading with the huge resources of depository banks.[26] It was now the era of the so-called "universal bank" that could both take deposits and make loans and trade the financial markets.

As repeal of Glass-Steagall approached, investment banking and commercial banking began a period of dramatic consolidation. After its repeal, large commercial banks became universal banks through acquisition of investment banks or by launching new trading opera-

---

23. Glass-Steagall Act, Pub. L. No. 73-66, 48 Stat. 162 (1933); *see also* Stephen P. Kenkel, *Banking and Securities Law*, 56 GEO. WASH. L. REV. 736, 736 (1998) ("In passing the Glass-Steagall Act (Glass-Steagall or the Act), Congress attempted to separate commercial and investment banking.").

24. *See* Kenkel, *supra* note 23, at 738 ("Section 16 [of the Glass-Steagall Act] restricts the securities activities of Federal Reserve member banks 'to purchasing and selling such securities and stock without recourse, solely upon the order, and for the account, of customers. . . .' Section 20 bars member banks from 'affiliating' with organizations 'engaged principally in the issue, flotation, underwriting, public sale, or distribution . . . of stocks, bonds, debentures, notes, or other securities.' Section 21 prohibits any organization engaged in the business of issuing, underwriting, selling, or distributing securities from accepting deposits. Finally, Section 32 prohibits the existence of interlocking management between member banks and firms engaged in the securities business.").

25. *See* Arthur E. Wilmarth, Jr., *The Transformation of the U.S. Financial Services Industry, 1975–2000: Competition, Consolidation, and Increased Risks*, 2002 U. ILL. L. REV. 215, 219 (2002) ("The GLB Act removed legal restrictions on affiliations between banks and securities firms by repealing two provisions of a 1933 statute popularly known as the 'Glass-Steagall Act.'").

26. *See id.* ("As a result of the GLB Act, banks can combine with securities firms and insurance companies to organize financial conglomerates under the structure of a 'financial holding company.'").

tions. Finally, during the crisis, Bear Stearns was scooped up by JP Morgan Chase and Bank of America absorbed Merrill Lynch, both with the direct involvement of the government. Weeks later, Morgan Stanley and Goldman Sachs converted to banks to steady themselves in the turmoil. Investment banks, other than regional and boutique investment banks, ceased to exist as a separate category from commercial banks.

The consolidation was widespread, resulting in a system of mega-banks, an oligopoly with vast market power.[27] While there is no available comprehensive study of oligopoly behavior in the trading markets, the banks engage in specific trading activities that are consistent with this behavior. And one need only consider the reports of the investigation into the manipulation of the London Inter-bank Offered Rate ("LIBOR") as an example of this behavior.[28]

Recent research by the Dallas Fed provides a window on this process.[29] The study observes that in 1970 the top five banks in terms of assets held 17% of aggregate bank assets. By 2010, the top five banks held 52% of aggregate assets.[30]

In parallel, the newly created derivatives market was exempted from all meaningful regulation in the Commodity Futures Modernization Act of 2000. Derivatives, dubbed by Warren Buffet as "financial instruments of mass destruction,"[31] were to be unregulated, creating a new, massively risky market overnight. Universal banks were handed an opportunity to dominate another trading market, one perfectly designed for their information and quantitative advantages—the four largest banks currently control 94% of the bank derivatives business in the United States.[32]

---

27. *See* Guriev & Kvasov, *supra* note 14 (analyzing the post Glass-Steagall oligopolistic behavior of banks in the debt and equity markets).

28. Reed Albergotti & Jean Eaglesham, *9 More Banks Subpoenaed over LIBOR*, WALL ST. J., Oct. 25, 2012, *available at* http://online.wsj.com/article/SB10001424052970203897404578079413742864842.html. The breadth of the investigation suggests that many large banks were aware of the practice and simply joined in to secure their share of the profits, a concerted effort by the large universal banks to manipulate a system affecting a broad swath of the public.

29. FED. RESERVE BANK OF DALL., 2011 ANNUAL REPORT: CHOOSING THE ROAD TO PROSPERITY (2011).

30. *Id.*

31. WARREN BUFFET, BERKSHIRE HATHAWAY ANNUAL REPORT: LETTER TO SHAREHOLDERS 15 (2002), *available at* http://www.berkshirehathaway.com/2002ar/2002ar.pdf.

32. OFFICE OF THE COMPTROLLER OF THE CURRENCY, QUARTERLY REPORT ON BANK TRADING AND DERIVATIVES ACTIVITIES—FIRST QUARTER 2012 (2012), *available at* http://occ.gov/topics/capital-markets/financial-markets/trading/derivatives/derivatives-quarterly-report.html.

II. UNDERSTANDING THE FINANCIAL MARKET SYSTEM

There is a common understanding that financial markets work by constantly finding an equilibrium that is defined by fundamental value. If the markets are large enough and if the information relevant to fundamental value is equally accessible, prices of securities and commodities will be forced toward the theoretically correct levels by competition among traders. Since all participants are equally motivated by greed, high transaction volume will quickly erase any price anomaly or inefficiency. Individual greed will be canceled out by universal greed.

The foundation for this way of thinking is the efficient market hypothesis,[33] whose most famous devotee is Alan Greenspan.[34] Under the hypothesis, no regulation is the best regulation. This is completely simpatico with the conservative ideology of the Reagan administration and its successors. It became the foundation of pre-financial crisis policy on financial regulation.

A corollary to the efficient market hypothesis is the representative agent model.[35] If price anomalies and inefficiencies are instantaneously eradicated, individual motivations and quirks of market structures are irrelevant.[36] The marketplace can be viewed as monolithic for purposes of analysis. Academic and other experts adopted this model and systematically de-emphasized factors other than the common interest in optimizing market value.[37] Having adopted the representative agent model, academic research ignored the potential for a system dominated by the exploitation of distortions often created by individual agents.

The representative agent model affects the approach to regulation dramatically. Certainly, bad behavior can be proscribed, such as fraud or market manipulation. But the representative agent model causes the rule makers to set standards that require greater culpability as a precondition to rule violation. If trading activity in pursuit of greed is good, no one activity should be singled out unless it is heinous in the extreme. This is an institutional reason that so few partic-

---

33. The hypothesis was originally articulated by University of Chicago economist Eugene Fama in 1970. Eugene Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, 25 J. FIN. 2 (1970).

34. Burton G. Malkiel, *The Efficient Market Hypothesis and Its Critics* (Princeton Univ., CEPS Working Paper No. 91, 2003).

35. RAGHURAM RAJAN, FAULT LINES 116 (2010).

36. *Id.*

37. *Id.*

ipants in the financial market debacle have been taken to task in legal proceedings.

History has not been kind to the efficient market hypothesis. As Greenspan famously observed, the events of the financial crisis could not have existed under such a market cosmology.[38] Instead, the market was riddled with misinformation and conflicts of interest.[39] Market participants did not act to preserve the integrity of the market as predicted by the hypothesis.[40] They acted to maximize short-term profits (and their bonuses) until the music finally stopped, precipitating bankruptcy or bailout.[41] Contrary to the representative agent model, the pursuit of narrow interests determined outcomes. Market participants were not monolithic at all. Their actual behavior resembled the lifeboat evacuation of the Titanic, with similar inefficient consequences.

The economic model that better fits with the events of the recent past is that of economist Hyman Minsky: "[O]nce we admit that institutions are man-made and at least in part the product of conscious decision, we must also face the effects of institutional arrangements on social results."[42] He argues "that almost all systems which are multidimensional, nonlinear, and time dependent are endogenously unstable."[43] In Minsky's view, periods of market stability are destabilizing and markets are inescapably incoherent. Markets are not predictably efficient. They are similar to a fluid system influenced by multiple forces that are difficult to either predict or measure. A price can be formed in a market, but contemplation of its perfection is a futile task.

Today's markets are far more consistent with Minsky's theories. While some information is broadly shared by market participants (the essential assumption of the efficient market hypothesis), the ever-increasing speed and capacity of information technology assures that the more powerful market participants will always enjoy an information advantage. Especially in modern, high-speed markets, the perception of facts is the driving force. Perceptions can be altered.

---

38. Bryan Knowlton & Michael M. Grynbaum, *Greenspan "Shocked" That Free Markets Are Flawed*, N.Y. TIMES, Oct. 23, 2008, *available at* http://www.nytimes.com/2008/10/23/business/worldbusiness/23iht-gspan.4.17206624.html.

39. PERMANENT SUBCOMM. ON INVESTIGATIONS, COMM. ON HOMELAND SECURITY & GOVERNMENTAL AFFAIRS, U.S. SENATE, MAJORITY AND MINORITY STAFF REPORT, WALL STREET AND THE FINANCIAL CRISIS: ANATOMY OF A FINANCIAL COLLAPSE (2011).

40. *Id.*

41. *Id.*

42. HYMAN MINSKY, STABILIZING AN UNSTABLE ECONOMY 9 (2008).

43. *Id.* at 11 n.9.

Altered perceptions of current circumstances created by information asymmetries, even for small periods of time, can introduce tremendous distortions.

This reality undermines the concept that markets behave like natural systems in equilibrium. Thus, other system concepts must explain them better. Fluid systems, described by chaos theory, are not good comparisons. The basic principle of chaos theory is that future outcomes cannot be predicted even if one has all relevant information about current conditions. In financial markets, cause and effect is relatively clear. It is, however, difficult to know when the effect may occur and whether it will happen at a single time or in stages.

There is a third possibility: self-organized criticality. This system construct was articulated by the physicist Per Bak and likens natural systems to tectonic plates.[44] It is entirely predictable that earthquakes will occur, but it appears to be impossible to forecast when and in what stages the forces built up by plate movements will be released. The example provided by Professor Bak is a child dribbling sand into a pile at the beach.[45] At first the size and shape of the pile is predictable.[46] As the pile accretes, however, it will reach a point of criticality that results in an avalanche.[47] How and when this occurs depends on the complexity of the system, that is to say the irregularity of the grains of sand that hold the pile together by friction as forces accumulate only to be released in the avalanche.[48] If the system were less complex, for instance if the particles were smooth like marbles, the force would not accumulate but would be released constantly.

Self-organized criticality aptly describes modern financial markets. The markets are characterized by booms (mortgages, dot-com stocks) and busts (financial crisis). This is very different from a system in constant equilibrium.

The application of Professor Bak's construct has extraordinary implications for evaluating the costs and benefits of regulation. It suggests how regulation should best be crafted to shape markets that efficiently price and deploy capita. Self-organized criticality derives from the complexity of the system. Thus reducing complexity has value in and of itself by mitigating the violence of the type of avalanche that we call a financial crisis. Instead of shying away from di-

---

44. PER BAK, HOW NATURE WORKS: THE SCIENCE OF SELF-ORGANIZED CRITICALITY 1–3 (1999).

45. *Id.* at 3.

46. *Id.*

47. *Id.*

48. *Id.*

rectly restricting transactions that create complexity, it should be a goal.

III. PERFORMANCE OF THE DEREGULATED MARKETS

If deregulated financial markets can best be described by self-organized criticality, complexity might still be a benefit to the economy if the markets are performing their social functions more efficiently than they might otherwise. One need only compare the increased financial sector compensation levels that accompanied deregulation with the persistent unemployment that follows recessions and income and wealth disparity to see that they are not.[49] In the referenced article, I describe connections between deregulation and profits and their share of the economy.

It is not simply that the financial sector has been extremely profitable, though it has been. During the thirty-five years of deregulation, the financial sector share of the economy has increased to unprecedented levels, growing from 3.8% to 8.2% of the GDP,[50] while the manufacturing and services sectors have become relatively smaller. The trend is not caused by exporting financial services to foreign nations.[51] Instead, the modern financial services sector is profiting by extracting the value that it delivers to the rest of the U.S. economy. This phenomenon is structural.

Aside from insurance and payment systems, the essential service of the financial sector is capital intermediation.[52] Sources of capital, such as savings and pension funds, must be matched up with users of capital who are financing productive activities.[53] The overwhelming value that is extracted by the financial sector, which relates to its profitability and share of GDP, is extracted from capital intermediation in various forms.[54] The efficiency of the matching systems is a function of fundamental capital cost and the cost of intermediation. The price

---

49. *See* Turbeville, *supra* note 1 (describing connections between deregulation and increased financial sector compensation levels, profits, and share of the economy and pointing out connections between these phenomena and persistent unemployment following recessions).

50. Thomas Philippon, *The Future of the Financial Industry*, STERN ON FIN. (Oct. 16, 2008), http://sternfinance.blogspot.com/2008/10/future-of-financial-industry-thomas.html.

51. Turbeville, *supra* note 1.

52. Thomas Philippon, *The Size of the U.S. Finance Industry: a Puzzle?* (Fed. Reserve Bank of N.Y., N.Y. Area Monetary Policy Workshop Paper, 2011), *available at* http://www.newyorkfed.org/research/conference/2011/NYAMP/Fed_Philippon_v1.pdf.

53. *Id.*

54. *Id.*

paid for intermediation must be rationally related to the service provided. In terms used by economists, the economic rent extracted for capital intermediation must be optimally small.

A recent study of intermediation costs by Thomas Philippon of New York University's Stern School of Management addressed the causes of the current large financial sector share of the economy.[55] Professor Philippon used the neoclassical growth model to examine financial intermediation in the United States over a 140-year period. He constructed an index that measures the unit cost of financial intermediation to show that the finance industry has become less efficient in providing intermediation services over time.[56] He summarized his findings as follows:

> [T]he finance cost index has increased since the mid-1970s. This is counter-intuitive. If anything, the development of information technologies (IT) over the past 40 years should have disproportionately increased efficiency in the finance industry. How is it possible for today's finance industry not to be significantly more efficient than the finance industry of John Pierpont Morgan a century ago? [The historic trends] present[] a puzzle for future research.[57]

Finance has obviously benefited from the IT revolution, which has dramatically lowered the cost of retail finance. Yet, even when accounting for all the financial assets created in the United States, the cost of intermediation appears to have increased. Why is the nonfinancial sector transferring so much income to the financial sector? Professor Philippon concludes that, mechanically, the reason is an enormous increase in trading.[58]

The study indicates that the cost of intermediation between the suppliers of capital and the productive consumers of capital has increased notwithstanding technology advances, sophisticated quantitative analysis, massive trading volume increases, and diversity in securities and derivatives markets.[59]

The only time over the 140-year period that Professor Philippon's financial intermediation cost index was comparable to the period of

---

55. Thomas Philippon, *Has the U.S. Finance Industry Become Less Efficient? On the Theory and Management of Financial Intermediation* (May 2012) (unpublished manuscript) [hereinafter Philippon, *Less Efficient?*], *available at* http://pages.stern.nyu.edu/~tphilipp/papers/Finance_Efficiency.pdf.

56. *Id.* at 16–17.

57. *Id.* at 15–17.

58. *Id.* at 22.

59. *Id.* at 21 tbl.11.

deregulation was the Great Depression.[60] High costs of intermediation make sense in the Great Depression when intermediation virtually ceased to exist—there was even a bank holiday for a period. But in the deregulation period, banks were profitable and the supply of financing was plentiful. Reasoning under the tenets of the efficient market hypothesis, the professor correctly concludes that this is absolutely counter-intuitive. But from the perspective of an observer of trading behavior and market evolution, his results make perfect sense.

Thinking of the financial markets in the context of self-organized criticality allows us to see behind Professor Philippon's reference to "an enormous increase in trading."[61] Complexity is the key to understanding the massive inefficiency of capital intermediation and increased volume is merely a byproduct of increased complexity.

By understanding how complexity fuels inefficiency we can better understand why the financial sector has incentives to increase complexity and why this leads to financial crises. Asymmetric information allows those market participants with better information to consistently earn more from trading than others.[62] Complexity allows more effective deployment of information asymmetry advantages since the available information is more difficult to discern and analyze.

The remainder of this Article examines real world applications of asymmetry. First, it looks at high frequency trading, which is driven by information technology. It then examines derivatives, which involve advanced quantitative analysis. These examples of the two forms of information asymmetry are by no means exclusive, but they are the most visible. Regulations that limit these two activities should not be seen as burdening marked efficiency. In contrast, such regulations should be assumed to increase efficiency of capital intermediation.

IV. HIGH FREQUENCY TRADING ("HFT")

If asked to describe the role of a financial market structure, a trading professional will virtually always say that it is to facilitate liquidity so that transactions can be accomplished efficiently and at a low cost. This makes sense because the compensation of traders and fund managers is related to transaction costs.

Properly used, efficient liquidity for a given transaction refers to market conditions needed so that initiation of the transaction will not, by itself, alter the best available transaction price. If, for exam-

60. *Id.*
61. *Id.* at 19–21.
62. *See supra* Part II.

ple, a large number of willing buyers are active in a market, a seller is more likely to receive the highest going price for a sale he posts to the marketplace seeking a bid. There is depth in the buying interest at the best going price that is sufficient to absorb the sale.

A large number of transactions taking place in a market generally means that transaction liquidity will be high and the seller is more likely to receive the price most recently bid to other sellers when his posted offer to sell is matched with a buyer. This assumes that high volume information traders, those whose strategy is to enter and exit the market quickly to take advantage of superior information, are consistently willing to transact at the going market prices and provide liquidity. In the modern markets, most information traders employ powerful computers driven by algorithms that can switch from providing liquidity to liquidity consumption instantaneously.

During any observation period, the price paid by capital consumers should be greater than the price received by capital investors. Traders must be compensated for intermediation, after all. It is often assumed that when the spread between the prices at which traders can buy and sell is as low as it can be, intermediation is efficient. Capital intermediation is relatively inefficient.[63] If the enormous increase in volume of trading in today's market represented by HFT is providing liquidity and lowering that spread, how can intermediation be inefficient?

High frequency trading has been defined as "fully automated trading strategies with very high trading volume and extremely short holding periods ranging from milliseconds to minutes and possibly hours."[64] At the speed of HFT transaction execution, no human decision-making is possible. The decisions are driven by algorithms that dictate the placement of orders and purchases or sales based on observed market conditions. A computer-driven trading operation enslaved to an algorithm is like a "trader-bot," intended to act just like a human trader but at high speed. As we shall see, this intent cannot be fulfilled because the flexibility of algorithms is limited.

High frequency trading is pervasive, especially in the equities markets where it has been estimated that it represents 73% of all volume.[65] It has changed fundamental characteristics of markets. At the end of World War II, the average holding period for stocks was four

63. *See* Philippon, *Less Efficient?*, *supra* note 55, at 16–17.

64. Frank Zhang, *High Frequency Trading, Stock Volatility, and Price Discovery* (Dec. 2010) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1691679.

65. R. IATI, TABB GRP., HIGH FREQUENCY TRADING TECHNOLOGY (2009).

years. By the turn of the millennium, it was eight months. By 2008, the average holding period declined to two months. And by 2011, it has been estimated that, at least for actively traded stocks, it had declined to twenty-two *seconds*.[66]

While there has been speculation that high frequency trading may have declined recently, a November 2012 study funded by the Commodity Futures Trading Commission focusing on HFT profitability in the equities futures market, an integral element of the equities market, finds that the percentage share of HFT in that sector has remained constant.[67] High frequency trading is the dominant form of information trader activity and a likely source of value extraction from the capital intermediation process.

The inquiry into the flash crash of May 2010 resulted in a study of market dynamics during that event by Andrei Kirilenko, the chief economist of the Commodity Futures Trading Commission, and others.[68] On May 6, 2010, the Dow Jones Average plunged over 1000 points in a matter of minutes.[69] This represented approximately $1 trillion of market value.[70] The culprit most cited was a mutual fund whose algorithms governing trading tactics triggered a $4 billion sale of equity instruments, on a day when the market was particularly shaky, with insufficient regard to price.[71] The market soon recovered, but the mayhem caused by the rapid moves was substantial.[72] Fortunately, the flash crash occurred in the early afternoon allowing time for the market to recover.

---

66. Paul Jay, *Interview with Michael Hudson*, REAL NEWS NETWORK (Jan. 1, 2011), http://www.therealnews.com/t2/index.php?option=com_content&task=view&id=31&Itemid=74&jumival=6000#.UQVySejaicY. Jay spoke with Professor Hudson, a professor of economics at the University of Missouri—Kansas City regarding this estimate. He reported that he had been told by a number of trading professionals that this was the case. He then examined data from the New York Stock Exchange and the Chicago Mercantile Exchange and compared shares of actively traded stock outstanding with activity in those shares on the exchanges. He concluded that the data were consistent with and reinforced the twenty-two second estimate. *Id.*

67. MATTHEW BARON, JONATHON BROGAARD & ANDREI KIRILENKO, NAT'L BUREAU OF ECON. RESEARCH, MARKET MICROSTRUCTURE MEETING, THE TRADING PROFITS OF HIGH FREQUENCY TRADERS (2012), *available at* http://conference.nber.org/confer/2012/MMf12/Baron_Brogaard_Kirilenko.pdf.

68. Andrei Kirilenko et al., *The Flash Crash: The Impact of High Frequency Trading on an Electronic Market* (May 26, 2011) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1686004.

69. David Easley et al., *The Microstructure of the 'Flash Crash,'* 37 J. PORTFOLIO MGMT. 118, 119 (2010).

70. *Id.*

71. Kirilenko et al., *supra* note 68, at 9.

72. *Id.* at 8.

The Kirilenko study targeted the role of algorithmic, high frequency trading as it interacted with the initial aberrant price move caused by the mutual fund.[73] The study contrasted behavior of HFTs and market makers, who, unlike information traders, profit in the market from the spread between purchase prices and sale prices rather than price movement.[74] The study showed that market makers tended to moderate the price moves because they did not instantaneously exit the markets, dumping inventories on the way out.[75] In contrast, HFTs, who seek to profit from price moves at high speed, exacerbated the flash crash by reacting, en masse, to price moves automatically as their systems responded to unusual market moves in ways dictated by their governing algorithms.[76] Kirilenko has described an "ecosystem of market participants" that interacted through algorithms and high-speed trading systems to create a serious stock market anomaly.[77]

The study demonstrates that there is a great distinction between volume and market liquidity. For example, algorithmic trading activity can amplify the price effect of a given market event.[78] Prior to an event, market participants misperceive the volume generated by the algorithmic traders as stabilizing liquidity. Yet the systems are rigged to exit the market and dump inventories at the worst possible time, in terms of stability. When an event occurs, the "stabilizing liquidity" converts instantaneously into trading that consumes massive amounts of liquidity. The perceived stabilizing liquidity is an illusion, an even more disruptive circumstance than if the illusory stabilizing liquidity had never existed at all.

The 2010 flash crash was a dramatic example of events that occur daily in the markets.[79] Mini-flash crashes distort markets repeatedly, causing prices to be unreliable and volatile. A recent study likewise found that high frequency trading is positively correlated with stock price volatility, after controlling for fundamental volatility and other

73. *Id.* at 2–3.

74. *Id.*

75. *Id.* at 27.

76. *Id.* at 36.

77. Andrei Kirilenko, Presentation to the Commodity Futures Trading Commission Technical Advisory Committee Roundtable (Oct. 12, 2010), *available at* http://capitolconnection.net/capcon/cftc/webcastarchive.htm#.

78. Kirilenko et al., *supra* note 68, at 36.

79. Graham Bowley, *The Flash Crash, in Miniature*, N.Y. TIMES, Nov. 8, 2010, at B1.

exogenous determinants.[80] The study identifies intra-day volatility that is actually a series of mini-flash crashes.[81]

The point is made eloquently by Eric Hunsader, the founder of Nanex, a high speed market date feed service:

> In summary, HFT algos reduce the value of resting orders [for example, a market maker's orders] and increase the value of how fast orders can be placed and cancelled. This results in the illusion of liquidity. We can't understand why this is allowed to continue, because at the core, it is pure manipulation.[82]

But what about the time periods in which HFTs are not exiting the marketplace precipitously? After all, a number of studies have concluded that HFT provides benefits.[83] These studies, however, focus on lower individual transaction costs. It is reasonable to find a benefit if the focus is narrowed to that level.

Another study focuses instead on the effect on the narrower class of market participants who merely intend to make investments.[84] This class of traders is referred to as value investors. The study also discusses the effect on HFT on market makers and information traders whose activities establish the spread between available purchase and sale prices.[85] The study characterizes the basic purpose of HFT as the use of speed to insert the HFT trader in between value investors and market makers.[86] The idea is to buy from or sell to a value investor and instantaneously sell to or buy from a market maker. The study concludes that the prices paid or received by value investors are adversely affected because liquidity-providing market makers adjust their price demands to compensate for the value extracted by HFTs.[87]

---

80. Zhang, *supra* note 64, at 1. In other words, all volatility caused by factors other than the HFT activity was filtered out and a strong correlation between HFT activity and the remaining volatility was found to recur frequently in each trading day.

81. *Id.*

82. *Enough Already!*, NANEX (Aug. 8, 2011), http://www.nanex.net/Research/EMini2/EMini2.html. The statement was made in an article referencing a particular algorithm known as "The Disruptor" that is specifically intended to disrupt the stabilizing order behavior of a liquidity provider.

83. *See, e.g.*, Terrence Hendershott & Ryan Riordan, *Algorithmic Trading and Information*, J. FIN. & QUANTITATIVE ANALYSIS (forthcoming) (manuscript at 22), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2001912.

84. Álvaro Cartea and José Penalva, *Where Is the Value in High Frequency Trading?* (Dec. 21, 2011) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1712765.

85. *Id.* at 3.

86. *Id.*

87. *Id.* at 1.

In other words, individual transaction costs are not increased but absolute price levels are altered to the detriment of the value investors, especially large institutional investors that transact large positions. Since value investors make their decision to invest based on return on funds invested, prices must adjust to compensate for the lost value.[88]

Investors don't merely buy a security and passively hold it. They invest in the market price for the security over time.[89] Further, they periodically replace the security by selling and acquiring another security. Valuation of an investment security and its successive replacements depends on the dynamic market price over time.[90] The investors mark the position representing the successive securities to current market prices daily.[91] Each day, they make the decision to hold the security or to replace it using these marks. When the investor allocates a sum of money to investment, he is investing in a series of daily transactions, including the decision to hold the security, that are affected by the changing fundamental value of the security and its successors, but also by the market activities of others that affect daily valuation via non-fundamental forces.[92]

When a business or government issues debt or equity, the price it receives is determined by the expected return required to induce investors to transact. If the expected return is uncertain, the price paid by the investor must be lower to compensate for uncertainty. The return may be uncertain because the business or government is risky from a fundamental value perspective, but it may also be uncertain because of market unreliability. For example, if the market lacks reliable liquidity for the security, the ongoing value is subject to price risk over and above fundamental value risk.[93] Liquidity may not be sufficiently high to generate purchases and sales at the fundamentally sound price over time.[94] The cause is likely to be information trader activity that consumes liquidity either consistently or periodically.[95]

---

88. *Id.* at 45.

89. *See, e.g.*, R. E. BAILEY, THE ECONOMICS OF FINANCIAL MARKETS 5 (2005).

90. *Id.* at 9.

91. *Id.* at 5.

92. *Id.* at 9.

93. Yakov Amihud & Haim Mendelson, *Stock and Bond Liquidity and Its Effect on Prices and Financial Policies*, 20 FIN. MARKETS & PORTFOLIO MGMT. 19, 21 (2006), *available at* http://asaprinceton.com/images/meeting/101711/Amihud_and_Mendelson_/amihud_and_mendelson2006.pdf.

94. *Id.*

95. *Id.*

Investment is not a structured process in which dollars can be tagged and traced through a linear series of transactions. There is a value extraction return premium attached to a security that is set by value investors to offset the cost of future market disruption. When a business or government raises money in the market, it pays that premium to the value investors. But the reason the value investors require the premium is that they know they must pay it back to information traders over time as a result of unreliable valuation prior to and at the time the position is liquidated or replaced. The premium ultimately is paid to HFTs as they extract value from the market on a daily basis.

The best empirical evidence of this is probably the financial sector's increased share of GDP and profits in the economy as a whole, as described above. No sector of the economy is better suited to benefit from advances in information technology and quantitative analysis. Furthermore, the rise of mutual funds and other aggregated investment funds has increased the scale of investment activities. The financial sector should be able to perform the intermediation function far more efficiently than in the past. On a transaction-by-transaction basis, these advances have resulted in lower transaction spreads as markets volumes have increased. Thus, the increased share of the economy represented by the financial sector, occurring simultaneously with transaction efficiencies that have had the effect of reducing the share, indicates that a substantial portion of market activity extracts value from the overall capital intermediation process.

In addition, the CFTC-sponsored study on profitability of HFTs cited above demonstrates persistent and substantial profitability for this large element of the trading market even though the risk taken on to earn this profit is very small.[96] Consistent profit, in excess of the going spread between purchase and sale prices with very little risk, means that the HFTs are extracting value based on structural, rather than fundamental, information. There is simply no other way that these conditions could co-exist. The HFTs are reaping the value extraction return premium that they caused to exist. And the persistent large profit that is disproportionate to risk means that the value extracted far exceeds liquidity benefits.

V. DERIVATIVES

Derivatives are the foundation for most of the financial engineering of the last thirty-five years. Valuation and risk metrics for deriva-

96. Baron, Brogaard & Kirilenko, *supra* note 67.

tives involve devilishly complex mathematics, but their structure is relatively straightforward.[97] One need not be a quant[98] to understand how they work and how they affect the efficiency of capital markets.[99]

A. *Basic Properties of Derivatives.*

As a threshold matter, it is essential to understand a few basic principles underlying derivatives. Financial institutions consistently describe derivatives in language that is designed to make them appear benign to customers and regulators.[100] Derivatives are characterized as financial products that reduce risk.[101] This facile description is parroted by academics and policymakers.[102] A new and more accurate description is badly needed.

The fundamental characteristic of a derivative is that it is a bilateral contract between two parties, requiring performance in the future.[103] It is not an asset like a share of stock or a barrel of oil. An existing derivative is not sold to another party. If a derivative counterparty wants to eliminate the derivative price risk from its book, it must enter into a contract that is the same, but take on the opposite obligations.[104] The new opposite-way derivative offsets the first derivative, but only if the party on the other side performs its obligations.[105]

The value to a counterparty of a derivative on any given day during its life involves two central properties of the contract: (1) The ex-

---

97. Adam R. Waldman, *OTC Derivatives & Systemic Risk: Innovative Finance or the Dance Into the Abyss?*, 43 AM. U. L. REV. 1023, 1027 (1994).

98. *See* Henry T.C. Hu, *Swaps, the Modern Process of Financial Innovation and the Vulnerability of a Regulatory Paradigm*, 138 U. PA. L. REV. 333, 338–39 (1989) (noting that financial institutions have become the domain of Ph.D.s, giving them the title "quants").

99. *See* Waldman, *supra* note 97, at 1027 (noting the conceptual simplicity of derivatives).

100. *See* Steve Denning, *Big Banks & Derivatives: Why Another Financial Crisis Is Inevitable*, FORBES (Jan. 8, 2013, 6:26 PM), http://www.forbes.com/sites/stevedenning/2013/01/08/five-years-after-the-financial-meltdown-the-water-is-still-full-of-big-sharks/ (noting that Wells Fargo hid derivatives behind benign labels).

101. *See* Norman Menachem Feder, *Deconstructing over-the-Counter Derivatives*, 2002 COLUM. BUS. L. REV. 677, 681 (describing derivatives as "financial products").

102. Bernard J. Karol, *An Overview of Derivatives as Risk Management Tools*, 1 STAN. J.L. BUS. & FIN. 195, 196 (1995); Peter S. Goodman, *Taking a Hard New Look at a Greenspan Legacy*, N.Y. TIMES, Oct. 9, 2008, at A1.

103. *See* Feder, *supra* note 101, at 691 (explaining the structure of derivatives).

104. Another method of eliminating risk is to "cash settle" the transaction. Michael Greenberger, *Overwhelming a Financial Regulatory Black Hole with Legislative Sunlight: Dodd-Frank's Attack on Systemic Economic Destabilization Caused by an Unregulated Multi-Trillion Dollar Derivatives Market*, 6 J. BUS. & TECH. L. 127, 130 (2011).

105. The party is still subject to the risk that other party's default will render them unable to meet their obligations. Feder, *supra* note 101, at 689.

pected financial value of the performance in the future by the other party to the contract; and (2) the likelihood that the required performance by the other party will not occur and that the expected financial value will not be realized. The values of these properties can be, and typically are, measured independently. Measurement of these values and how they interact, even for a simply structured derivative, is a complex task.[106]

A swap is a type of derivative contract to exchange one value for another, the value most often being the price of a security or other asset (the referenced asset).[107] Swaps are structured on hundreds of different prices, including prices of equity shares, currencies, energy and agricultural commodities, precious and commercial metals, and debt.[108]

The basic building blocks of swaps are forward prices.[109] A forward price, as of any date, is the expected price of a referenced asset on a specified date in the future. The expectation is represented by other recent contracts on the same forward price, an agreed index, or sometimes even an estimate by one of the parties to the contract. For example, a simple oil price swap is based on the forward price assuming a future delivery date and location and a quantity. A common contract might refer to June delivery of 100 barrels of crude oil at Cushing, Oklahoma, a major distribution pipeline hub.

One of the values to be exchanged in the future may be fixed and determined at the inception of the contract.[110] In this case, the performance required of one of the parties ("Party *A*") is payment of a set dollar amount.[111] This payment is calculated in our oil swap example as the market value, at the inception of the swap, of 100 barrels of oil to be delivered in June at Cushing, Oklahoma. If the current forward price for a barrel of oil to be delivered in June is $100, a swap on 100 barrels of oil would require a fixed payment by Party *A* in June of $10,000.

At least one of the values to be paid under a swap is an amount based on a price to be determined definitively in the future.[112] In our

106. Sanjeev Arora et al., *Computational Complexity and Information Asymmetry in Financial Products* 1 (Princeton Univ. Dep't. of Comp. Sci., Working Paper, 2012), *available at* http://www.cs.princeton.edu/~rongge/derivative.pdf.

107. Greenberger, *supra* note 104, at 132.

108. Feder, *supra* note 101, at 681; Karol, *supra* note 102, at 200–04.

109. Dan Awrey, *Regulating Financial Innovation: A More Principles-Based Alternative?*, 5 BROOK. J. CORP. FIN & COM. L. 273, 297 (2011).

110. Greenberger, *supra* note 104, at 132.

111. *Id.*

112. *See* Feder, *supra* note 101, at 691–706 (giving an overview of derivatives).

example, the other party to the oil price swap ("Party *B*") is required to pay an amount in June equal to the *then current* price of 100 barrels of oil delivered at Cushing. Thus, the values exchanged on performance of the contract are the forward price for oil delivered in June at inception and the actual price for oil delivered in June.

For Party *A*, the realized financial value of the future performance of Party *B* depends on the delivery price in June. If that price is higher than the forward price at inception of the swap (Party *A*'s fixed payment obligation), the swap will have a positive value for Party *A* equal to the excess. If it is lower, the swap will have a negative value to Party *A* equal to the amount the price has gone down since inception. Party *B*'s value in each of these cases will be the inverse. To put numbers on it, remember the example in which the June forward price of oil was $100 per barrel at the inception of the swap, resulting in a fixed payment by Party *A* of $10,000. If the price per barrel is $120 on the performance date in June, Party *B* will be required to pay $12,000. After netting out the fixed payment by Party *A*, Party *B* is out $2000, which is also the benefit to Party *A*. But if the price is $80 per barrel, Party *B* will receive, and Party *A* will pay, $2000 on a net basis.

On each date prior to the definitive determination of the floating payment (Party *B*'s performance obligation), the accrued value of the swap is based on the forward price on that date. Thus, the dynamic value of any derivative is dependent on the movement of a referenced forward price over a specified time period.

Assuming that the fixed payment (made by Party *A*) is accurately based on the June Cushing forward price at inception, the swap initially has no intrinsic financial value. At inception, the expected amount to be paid by Party *B* is the same as the fixed amount to be paid by Party *A*. As soon as June Cushing forward price changes, however, value accrues. If the price increases the very next day after inception by $1, Party *A* accrues that positive value; if it decreases by $1, the accrued positive value is Party *B*'s. The counterparty in each case accrues an equal negative value.

But, in contrast with this accrued value, the *realized* financial value is not actually known until performance is completed. The swap is a contract that has realized value only if the other party performs.[113] It is not an asset that can be converted to cash by selling it on a date chosen by the owner. If Party *A* accrues a $1 value for each barrel, that value is at risk if Party *B* goes bankrupt or otherwise fails to per-

113. If, for example, Party *B* defaults and is unable to meet his obligations to pay, then Party *A* realizes nothing.

form. Party *A* is exposed to the credit of Party *B*. Party *A* can replace the swap immediately if Party *B* goes bankrupt. But the replacement swap will have a new inception date and, on that date, the June forward price is $1 higher in our example. Therefore, the fixed payment is based on a $1 higher price for 100 barrels. The consequence to Party *A* when June comes around is $100 worse.

This consequence of insolvency of Party *B* prior to performance is exactly the same as if Party *A* had loaned $100 to Party *B* and then Party *B* went bust. Party *A* has extended credit to Party *B*. That is exactly how sophisticated market participants, like banks, view swaps.

This view illustrates the second embedded property of a swap that determines its value. At any time the swap has a financial value based on the current forward price; but it also involves an extension of credit that has a separate value or cost.[114] In a loan, this value and cost is expressed as an interest rate. In a derivative, the credit extension has a parallel value or cost, but it its expression is obscured because it is embedded in the pricing of the swap. For example, the floating leg of the swap may be the June Cushing price less fifty cents. This means that the swap has intrinsic value that compensates Party *A* for the potential extension of credit to Party *B*. Sophisticated derivatives counterparties price in potential credit extension using complex statistics, especially if they have market power to demand it.

B. Valuation of Cost of Derivatives to Capital Intermediation.

Capital intermediation is the process of matching available investment funds with productive uses of capital investment.[115] Mismatches can be based on supply and demand. But, there are other mismatches that arise from particular needs of investors and businesses and governments.

Derivatives are an integral part of the intermediation provided by trading markets.[116] In theory, they reconcile mismatches between capital sources and uses, typically interest rate, creditworthiness, and currency differentials. They act just like bank capital reserves in the commercial bank intermediation model.

For example, a company or government can enter into a derivative, contract with a bank to synthetically convert the form of an obligation into another form. Using derivatives, an investor who seeks a

114. *See* Feder, *supra* note 101, at 687 (explaining that credit risk exposure is an inevitable consequence of entering into a derivative).

115. Turbeville, *supra* note 1.

116. *Id.*

ten-year, fixed rate bond denominated in Euros can be a source of funds for a company seeking a floating rate loan denominated in U.S. dollars. The company can synthetically convert the bond that the investor wants to buy into funding based on terms that the company wants to procure.

In our example, the company enters into a swap contract with a bank. Under the contract, the company receives payments equal to its fixed rate interest payment obligation and it pays the bank an amount equal to the interest obligation that it would have had if the interest obligation had been at a floating rate. The investor is paid its fixed rate interest indirectly by the bank and the company pays floating amounts to the bank. The company's obligations denominated in Euros are similarly swapped with the bank for a like obligation denominated in U.S. dollars. Capital intermediation has been achieved and the mismatches have been reconciled. The question is whether it has been achieved more or less efficiently than would be the case if alternatives had been used.

Derivatives do not eliminate risk.[117] They are contracts that exchange one set of future consequences from a price change for another, assuming the other party performs.[118] Picture a business whose profit and loss during a period in the future depends on price movements of a commodity or security. To avoid the consequences of an adverse price move, the business could establish a reserve from borrowings or earnings. Alternatively, it could enter into a swap that (assuming performance by the counterparty) fixes the consequence of this price exposure at the current price level.

The distinctions between these alternatives—capital reserves and derivatives—should drive the decision between these two methods of managing the risk of price movement. It is similar to the decision between buying insurance and self-insuring. The structural differences between the two alternatives have been found to be relatively small (though the cited study fails to consider important risks that are embedded in derivatives).[119] If a reserve is used, the business must pay for the capital to fund it. If a swap is used, the company pays the value of a beneficial price move if it occurs plus embedded charges. If a reserve is used, the risk is that an adverse price move has consequenc-

117. *Id.* at 683.

118. *See supra* Part VI.A.

119. Andrea Gamba & Alexander J. Triantis, *Corporate Risk Management: Integrating Liquidity, Hedging, and Operating Policies* 2 (Robert H. Smith School, Research Paper No. RHS 06-106, 2011), *available at* http://www.rhsmith.umd.edu/cfp/pdfs_docs/papers/Triantis CorporateRiskMgmt.pdf.

es beyond the reserve. If a swap is used, the basic risk is that the counterparty fails to perform, but swaps involve many other risks as well.[120]

The relative profitability of derivatives is legendary. In 2007, I made a proposal to the head of an energy division of one of the largest banks in the world whose responsibilities included lending to energy companies. The bank had allocated a large amount of credit capacity for certain companies to its derivatives desk. I suggested that the bank transfer some of the credit capacity for certain energy companies to the lending group. This would enable the companies to use loan proceeds to collateralize derivatives credit exposures to the bank rather than merely grow exposures organically by transacting derivatives. The collateralization would be done through a system my company had developed to increase the efficiency of derivatives credit risk management. If the bank charged uniformly for the extension of credit, it would be indifferent as to the allocation of the credit capacity since the exposure to the bank would be the same. Under my proposal, the exposure would be under a loan and the derivatives exposure would be fully collateralized with treasury securities. Therefore, the aggregate exposure would be the same. It appeared to be a matter of form rather than substance, but the customer would benefit from the increased risk management efficiency my company would provide.

The derivatives desk refused to re-allocate the credit capacity because the bank explained that its profit from credit extension embedded in derivatives was *ten times* the profit that could be earned from making a loan. This experience confirmed what was commonly understood to be true.[121] Because the pricing of derivatives was so complex, customers almost never understood how much a bank charged for entering into the derivative.[122] Assuming that this price disparity is accurate, the price for credit embedded in a derivative would be ten times the price for a loan to fund a reserve. The amount extracted by the financial sector is ten times more than the same lending transaction.

---

120. *See* Feder, *supra* note 101, at 721–31 (explaining the risks inherent in swaps).

121. *Frontline, Money, Power and Wall Street, Part One* (PBS television broadcast Apr. 24, 2012), *available at* http://video.pbs.org/video/2226666502. The transcript for this broadcast is available at http://www.pbs.org/wgbh/pages/frontline/business-economy-financial-crisis/money-power-wall-street/transcript-19/.

122. *See* Louise Story, *A Secretive Banking Elite Rules Derivatives Trading*, N.Y. TIMES, Dec. 12, 2010, at A1 (explaining that customers are unaware of how much they are being charged for entering into a derivative agreement).

This practice constitutes a massive distortion of the credit markets. In an efficient market, credit is priced similarly regardless of how it is deployed. Credit capacity is finite. The companies and governments seeking capital financing are constrained when available capacity is depleted. The large disparity in the profitability of lending and derivatives credit extension means that the businesses and governments will be prevented from tapping into capital lending sources or will pay more for scarcer capacity. It would be different if the credit extension were ten times more valuable to the company or government than a loan. Academic research suggests, however, that this is not the case, especially when the research is read in the context of the practical use of derivatives in the marketplace.[123]

Despite their enormous relative cost, companies and governments still use derivatives to hedge risks. Sometimes derivatives are used to obscure the truth from others. That was the case with the government of Greece, which entered in to an off-market swap with credit extended by the bank at inception to meet the European Union debt ratios. But far more often, the complexity of derivatives hides their true cost from the companies and governments that use them. A study by Gamba & Triantis, while the most comprehensive available, nonetheless fails to master the incredible complexity of even a relatively simple swap transaction,[124] illustrating clearly that derivatives are difficult to value. More complex derivatives are even more challenging as additional risks compound the valuation problem. It is completely unrealistic to believe that participants in the derivatives markets accurately value the transactions that they enter into. A recent study describes this problem:

> The practical downside of using derivatives is that they are *complex* assets that are difficult to price. Since their values depend on complex interaction of numerous attributes, the issuer can easily tamper derivatives without anybody being able to detect it within a reasonable amount of time. Studies suggest that valuations for a given product by different sophisticated investment banks can be easily 17% apart and that even a single bank's evaluations of different tranches of the same derivative may be mutually inconsistent.[125]

123. Gamba & Triantis, *supra* note 119, at 4.
124. *Id.*
125. Arora et al., *supra* note 106 (citations omitted).

Under these conditions, it is clear that the very complexity of valuation of derivatives constitutes a major inefficiency in the intermediation of capital sources and capital uses.

The inadequate valuation of derivatives imposes inefficiencies and costs on the intermediation of capital in the financial markets and contributes to the increased cost of intermediation observed by Professor Philippon.[126] Sophisticated banks understand valuation far better than their customers, a major asymmetry of information.[127] In this way, the financial sector extracts value from the financial markets through derivatives prices that are rarely understood by customers.

VI. CONCLUSION

The evaluation of the costs and benefits of limiting human activity through regulation is an ongoing and important discussion. When the activity is commercial, the discourse is subject to a bias that restricting activity reduces productivity in the economy to the detriment of wealth creation. When a politician says that a rule is "unduly burdensome," he means that the discernible benefits do not outweigh the wealth that would be created if the rule did not exist. Today, the most sophisticated and powerful articulation of this concerns the regulation of the financial markets through financial markets reform under existing and potential future legislation.

This debate suffers from preconceptions that rely on unfounded economic theory and misconceptions based on the sheer complexity of modern finance. The first challenge is to establish how to measure the value of a market structure that is defined by a given set of regulations. The best measure looks to the structure's value to society. The primary function of financial markets in this context is the intermediation of capital investment. If capital intermediation is efficient in a construct, the rent extracted by the financial sector for facilitating that function should be proportionate to the value added.

The evidence shows that this is not the case. The predominant approach that emphasizes the cost of individual financial transactions is clearly inadequate. The leading theory implicitly assumes that capital intermediation can be viewed as a linear path of a unit of investment from an investor to a productive use. In this view, the costs of individual transactions along the way determine the cost of the capital intermediation process.

126. Philippon, *Less Efficient?*, *supra* note 55.
127. Story, *supra* note 122.

This is naïve or intentionally misleading. Modern capital and derivatives markets are exceedingly complex and involve multiple methods for extraction of value by the financial sector that must be paid for by the productive economy. The amount extracted is demonstrably far higher than historic data or reasoned analysis suggests could possibly be reasonable. The rents extracted by the financial sector for intermediating capital investment are inefficiently high.

The cause of this inefficiency is asymmetric information. The financial sector is incentivized to promote complexity to maximize the value of information asymmetry. A byproduct is increased risk of catastrophic financial crises.

The costs and benefits of financial market regulation should not be biased by the assumption that the restriction of activities is a burden that must be offset by legitimate benefits such as safety and soundness of the system or even price transparency. Under this approach, regulations that reduce complexity or market power do not prohibit the financial system from providing its essential social function, but very likely enhance the efficiency of that process.