American Economic Review: Papers & Proceedings 2013, 103(3): 393–397
http://dx.doi.org/10.1257/aer.103.3.393

SPECULATION, INSURANCE, AND FINANCIAL REGULATION‡

Benefit-Cost Analysis for Financial Regulation†

By Eric Posner and E. Glen Weyl*

The Dodd-Frank Wall Street Reform Act calls for extensive rule-making for financial regulation, with details left up to relevant enforcement agencies. To help ensure this wide discretion is not abused, regulators should be required to use Benefit-Cost Analysis (BCA). While BCA has been applied extensively in environmental, health, and safety (EHS) regulation and antitrust analysis, it has little history in financial regulation (Whitehead 2012).

The basis of BCA in the former areas is an extensive literature in economics establishing a framework (Harberger 1971) and helping clarify key parameters (Viscusi and Aldy 2003). By contrast, we are not aware of any analogous literature in financial economics. Most work in asset pricing is concerned with informational, rather than allocative, efficiency (whether prices are predictable rather than whether welfare is maximized). Normative work in corporate finance focuses on a relatively narrow set (viz. single bank prudential regulation) of issues compared with those relevant to a regulatory authority and on qualitative mechanisms rather than the quantitative trade-offs at the heart of BCA.

In this paper, we make a modest start toward filling this gap. When an agency proposes a regulation, it should compare the compliance costs and the benefits. The former will usually be straightforward to calculate, and so in each of the first three sections of this paper, we examine how three different types of regulatory benefits can be quantified: avoiding systemic crises, solving informational externalities, and reducing gambling. We conclude in Section IV by addressing when BCA should be applied to regulation or to actions (such as introducing a new product) by a firm.

I. Cost of a Statistical Crisis

The central trade-off in much EHS regulation is between costs incurred with certainty and a reduction in the probability of extreme harms to human life or health. The value of a statistical life (VSL), the willingness of individuals to pay to reduce the probability of death, has become perhaps *the* central economic parameter used to evaluate EHS regulations.

Broadly, financial regulation has a similar structure. Stricter regulations, such as tighter capital adequacy standards or limits on the breadth of activities institutions can undertake, slow the circulation of credit and liquidity. However, they also tend, at least when properly designed, to reduce the chance of both individual bank failures and systemic crises. While these costs are perceived by the economy with very high probability, and are thus analogous to the costs of EHS regulations, the latter benefits mostly reduce the probability of a catastrophic negative outcome. Unfortunately, a parameter for translating such a reduced probability of a crisis into a dollar value with certainty, call it the *cost of a statistical crisis* (CSC), has received far less attention than has VSL.

In fact, we are not aware of any work that has proposed a value for this parameter. Various studies have considered and come to conflicting views about the social cost of economic fluctuations more broadly (Lucas 1987; Chauvin, Laibson, and Mollerstrom 2011), but Reinhart

‡*Discussants:* Patrick Bolton, Columbia University; Hayne Leland, University of California-Berkeley; Darrell Duffie, Stanford University; Markus K. Brunnermeier, Princeton University.

* Posner: University of Chicago Law School, 1111 60th Street, Room 420, Chicago, IL 60637 (e-mail: eric_posner@law.uchicgao.edu); Weyl: Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637 (e-mail: weyl@uchicago.edu).

† To view additional materials, and author disclosure statement(s),visit the article page at http://dx.doi.org/10.1257/aer.103.3.393.

and Rogoff (2009) document that the economic consequences of financial crises typically differ dramatically from those of other economic cycles. Estimates of the CSC combining the methodologies of these literatures are crucial for BCA.

Research proposing a parameter value will face many of the same problems that the VSL literature confronted. Despite these, we believe that, by the same logic applied in the use of VSL, the waste associated with having no commonly used CSC value is typically greater than that associated with having a quite inaccurate number. Agreement on a figure in the range 150 billion to 3 trillion dollars (viz. a crisis cost between 1 percent and 20 percent of US GDP of approximately 15 trillion dollars) would seem relatively easy to reach given the widely respected estimates of Reinhart and Rogoff. We would advocate a figure in the 1–2 trillion dollar range. On its own, implementing such a standard would eliminate many crossed decisions between different agencies and regulators that lead society to violate transitivity. While decisions made outside of this range might seem rare, experience with EHS regulations indicate that in the absence of a numerical benchmark extreme waste in both directions is not only possible but common (Hahn 2004). The one-and-a-half order-of-magnitude range casually suggested here is already only about twice as broad in logarithmic terms as consensus views about VSL (Viscusi and Aldy 2003) and the narrower range we are sympathetic to is no broader in logarithmic terms.

Agencies will also need to estimate the magnitude of risk reduction associated with different regulatory options. This will be at least as challenging as determining a number for CSC; however, if agencies are forced to make explicit their implicit estimates, it will stimulate research and criticism, ultimately improving accuracy.

II. The Allocative Value of Price Discovery

Much of the asset pricing literature concerns *informational efficiency* of market prices, that is, the tendency, or not, of financial markets to bring asset prices into line with the risk-adjusted present-discounted value of the cash flows the asset will generate. Many, especially nonbank, financial regulations are either promulgated with the goal of aiding such informational efficiency (e.g., limits on automated trading and transparency mandates) or criticized by opponents for impairing the informational efficiency of markets (e.g., punitive regulation of short-selling and the "Tobin" tax). Hirshleifer (1971) famously argued that the private supply of information to prices will typically not agree with the socially optimal supply.

On the one hand, some of the value of information is captured by the other side of a trade as the price moves to its new equilibrium, leading investments in informational trading to be insufficient. On the other hand, traders have an incentive to be the first to incorporate a piece of information into market prices, even if the social value of this acceleration is small, leading to excessive investments in accelerating the pace of adjustments. As Mankiw and Whinston (1986) argue, if an activity is over- (under)rewarded relative to its social value, it will stimulate excessive (insufficient) expenditures on entry and information acquisition costs.

Yet we know of no quantitative model articulating when trading is over- or undersupplied and to what extent. Suppose that individuals in a market believe, because of a lack of information or confusion, for a length of time T that the value of an asset is p (and, thus, it trades for this price) even though the true value of the asset is $p^\star$. What is the social loss from this mispricing?

For a loss to emerge, some decision of real economic consequence must depend on this price signal. While the price theory logic below is quite general, we present, for the sake of definiteness, a simple model here in which the quantity of the asset that exists depends on its price. Let $q(p)$ be the quantity of the assets that exists when, during the period of length T, the market price of the asset is p. If supply is linear,

$$q(p) = q^\star\left(1 + \frac{T\epsilon^\star(p - p^\star)}{p^\star}\right),$$

where $q^\star$ is the equilibrium quantity when price is $p^\star$, and $T\epsilon^\star$ is the elasticity of the asset's supply with respect to a change in price per unit time over the period when $p = p^\star$. It seems reasonable to assume this elasticity is proportional to time if the arrival rate of opportunities to create or dispose of the assets is close to constant.

Assuming that the market for producing the asset is efficient and supply is linear, the welfare of asset suppliers over this period is

$$\frac{q^\star}{p^\star}\int_{\hat{p}=0}^{p} p^\star + T\epsilon^\star(\hat{p}-p^\star)\,d\hat{p}$$
$$= q^\star\left[(1 - T\epsilon^\star)\,p + \frac{T\epsilon^\star p^2}{2p^\star}\right],$$

while, for simplicity assuming all asset purchasers value the asset at its true price $p^\star$, their welfare is

$$(p^\star - p)\,q(\hat{p}) = q^\star\left[p^\star - p - \frac{T\varepsilon^\star(p^\star - p)^2}{p^\star}\right].$$

Adding these and simplifying, total welfare is

$$q^\star\left[p^\star\left(1 - \frac{T\epsilon^\star}{2}\right) - \frac{T\epsilon^\star}{2}(p - p^\star)^2\right].$$

This is maximized at $p = p^\star$, so the loss from mispricing is the Harberger (1964) triangle

$$\text{(1)} \qquad \frac{T\epsilon^\star q^\star (p - p^\star)^2}{2}.$$

While this expression is quite general and independent of details of the market's operation, the private profit to be made from correcting the mispricing by buying or shorting the asset until the price gap is closed depends on the microstructure of the market. The more liquid a market is (the less prices adjust to large purchases or sales) the greater will be the private profit. In any case, under a variety of models (monopolistic, competitive with trade limits, etc.), the profit is *linearly* proportional to the price difference $|p - p^\star|$ and to the size of the market, which if $|p - p^\star|$ is not too large, is approximately $q^\star$. We can therefore represent the private profit as

$$\text{(2)} \qquad lq^\star|p - p^\star|.$$

A first pass at quantifying the distortions of private incentives obtains from taking the ratio of expressions (1) and (2):

$$\text{(3)} \qquad \frac{T\epsilon^\star|p - p^\star|}{2l}.$$

A few qualitative results emerge immediately from this analysis. First, corrections to small mispricings will be socially overincentivized and, thus, oversupplied relative to corrections of large mispricings. Second, mispricings that persist for a long period will go undercorrected relative to those that are short-lived. Third, mispricing of assets whose supply responds elastically to market prices are undercorrected relative to assets in fairly fixed supply.[1] Finally, arbitrage of mispricings in liquid markets will be oversupplied relative to those in less liquid markets.

The first two points provide a simple quantification of the common intuition that while expenditures on the acceleration of high-speed trading on many small bets is largely waste, arbitrage activity to close long-standing bubbles is likely to be undersupplied relative to the social optimum. The third point expresses a different common intuition that improving the pricing of markets whose prices have little impact on real economic activity (e.g., certain derivative securities) has little value. The fourth point is perhaps best seen as a corrective to the first three: if policy interventions are being considered that impact the liquidity of markets they will be beneficial or harmful to the extent that they target liquidity at markets where otherwise correction of mispricing would be undersupplied or oversupplied, respectively. This logic highlights the attractiveness of some commonly advocated policies, such as a small Tobin tax on transactions that would filter out "small" arbitrages, as well as proposals to reduce the frequency with which trades can be made, while cautioning against others, such as high collateral requirements for long-term short positions that might reduce the incentive of market participants to pop bubbles.

By the same logic as in the derivation above, the social loss from failing to correctly incentivize the correction of mispricing is proportional to the square of the difference between expressions (1) and (2) and to the elasticity of arbitrage

[1] This result is, perhaps, a bit overstated, as is the temporal dimension, because the more elastic is economic activity over a period the easier it will be to profit on the differences in prices. However, we follow most of the finance literature (Kyle 1985) in believing that real economic activity elasticity is a small part of the total liquidity of an asset, and, thus, that this dampens our result only slightly.

activity itself with respect to the profit it yields. Thus, the costs and benefits of policies affecting market liquidity or any other factor facilitating or inhibiting price discovery may be quantified by measuring the parameters above and the elasticity of arbitrage activity with respect to its rewards.

III. Gambling versus Insurance

Along with their informational role, perhaps the most commonly touted virtue of financial markets is the efficiency with which they allocate risks to those most able to bear them, providing insurance against shocks individuals face. On the other hand, if different investors have, or act as if they have, different priors over risks in the economy but have similar wealth, risk-aversion and exposure to risks, they will tend to engage in bets against one another that increase risk (or allocate it less efficiently under any of their priors) but allow both to believe they are benefiting at the other's expense (Weyl 2007). To the extent that financial interventions or regulations aid (e.g., position limits or asset purchases that prevent market collapse) or limit (e.g., prohibitions or high capital charges for innovative assets) market completion, regulators must account for the benefits and costs accompanying such new assets.

The simplest, though controversial, case in which market completing may be harmful is when different individuals simply have different beliefs. In this case, Brunnermeier, Simsek, and Xiong (2012) propose a criterion for determining if one or more transactions is beneficial or harmful. They argue that a transaction should be deemed inefficient if *under any single belief that is a convex combination of the beliefs held by the agents engaged in the trades* the transaction is Pareto-dominated by the transaction not taking place and instead some transfers being made among the agents. That is, all agents or anyone with a belief between theirs agree that the transaction is wealth-destroying in aggregate.

One natural way to make this principle quantitative is to adopt the least interventionist evaluation of transactions possible. This would count the social cost associated with a transaction that agents desire to take place as the smallest outside subsidy that would have to be given so that, under some convex-combination belief, the transaction would not be dominated by any set of transfers. Similarly, beneficial privately desired trades could be treated in an equally generous way, as generating gains equal to the largest tax that could be imposed on the trade such that there is some convex-combination belief under which it is not dominated by some set of transfers. Essentially this asks what is the most libertarian belief-consistent interpretation of welfare from the trades consistent with the actions of the agents and other available information. An alternative standard would be to estimate the true distribution of outcomes (viz. true belief) from available evidence, including the behavior of agents, and take the average welfare loss or gain from the trade.

In many cases, transactions that appear to be driven by differences in beliefs rather than in their utility functions or endowments do not originate, fundamentally, in belief differences but instead in the informational setting in which they operate. For example, suppose an individual invests her money with an active manager but is unaware of the full set of vehicles in which that manager is able to invest. If she observes only a subset of the dimensions of the products in which the investor places her money (e.g., average annual return and a coarse summary risk such as a credit rating), her manager will have an incentive to invest in products that perform well along these dimensions even at the cost of performing poorly along the dimensions that are unobservable to the investor (Holmström and Milgrom 1991). The agent, combined with her manager, will thus act as if they are "optimistic" about their performance of products that do well along these observable dimensions but poorly along dimensions unobservable to the investor. If different agents are observable to different investors, who have different managers, opening markets between these investors may be harmful as it leads each investor's manager to exploit her imperfect information more effectively. Similar arguments apply for imperfectly informed principals of other sorts, such as tax authorities or capital regulators.

The welfare analytics of this more indirect channel for gambling are similar to those when there are real differences in prior beliefs. In both cases, different reduced-form investors act as if they had different beliefs and engage in trades that are wealth-destroying under any convex combination of their perceived beliefs. The only significant difference is that the principal-agent scenario might call for stronger intervention both because the libertarian arguments are less compelling in this case and because it might be even more appropriate to ignore the "beliefs" of the two investors as these are clearly distorted by the imperfections in their information. In both cases, of course, the standard insurance benefits of market completion must be weighed against the harms from gambling in judging the net value of market completion. In Posner and Weyl (forthcoming) we provide informal examples of this calculus as applied to a range of new assets created over the past two centuries.

IV. Conclusion

In this paper we propose three principles for the quantitative evaluations of normative trade-offs in the regulation of financial markets. While we anticipate that in the near future such analysis will be used primarily to evaluate the costs and benefits of restrictions on the operations of markets, in Posner and Weyl (forthcoming) we advocate an alternative baseline for new derivative securities. In particular, we argue that a BCA should be applied to the introduction of new products into markets by private participants. Whether this more precautionary approach (products or practices are disallowed until they pass a BCA) or the more traditional, libertarian approach (regulations restricting products or practices must pass a BCA) is appropriate depends on the circumstances and should, itself, likely be subject to a BCA.

The importance of developing methods for benefit-cost analysis for financial regulation can scarcely be overstated. In recent years, courts have awakened to the fact that many such regulations lack a sound economic basis and have started blocking them (Trindle 2012). Agencies are scrambling to develop reliable methods; we hope future research will come to their aid.

REFERENCES

Brunnermeier, Markus K., Alp Simsek, and Wei Xiong. 2012. "A Welfare Criterion for Models with Distorted Beliefs." http://scholar.princeton.edu/markus/files/welfare10.pdf (accessed February 12, 2013).

Chauvin, Kyle, David Laibson, and Johanna Mollerstrom. 2011. "Asset Bubbles and the Cost of Economic Fluctuations." *Journal of Money, Credit, and Banking* 43 (S1): 233–60.

Hahn, Robert W. 2004. "The Economic Analysis of Regulation: A Response to the Critics." *University of Chicago Law Review* 71 (3): 1021–54.

Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review* 54 (3): 58–76.

Harberger, Arnold C. 1971. "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay." *Journal of Economic Literature* 9 (3): 785–97.

Hirshleifer, Jack. 1971. "The Private and Social Value of Information and the Reward to Inventive Activity." *American Economic Review* 61 (4): 561–74.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7 (S): 24–52.

Kyle, Albert S. 1985. "Continuous Auctions and Insider Trading." *Econometrica* 53 (6): 1315–35.

Lucas, Robert E., Jr. 1987. *Models of Business Cycles*. New York: Blackwell.

Mankiw, N. Gregory, and Michael D. Whinston. 1986. "Free Entry and Social Inefficiency." *RAND Journal of Economics* 17 (1): 48–58.

Posner, Eric, and E. Glen Weyl. Forthcoming. "An FDA for Financial Innovation: Applying the Insurable Interest Doctrine to 21st Century Financial Markets." *Northwestern University Law Review* 107 (3).

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton, NJ: Princeton University Press.

Trindle, Jamila. 2012. "Another Scalia Vexes Regulators." *Wall Street Journal*, October 2.

Viscusi, W. Kip, and Joseph E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World." *Journal of Risk and Uncertainty* 27 (1): 5–76.

Weyl, E. Glen. 2007. "Is Arbitrage Socially Beneficial?" http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1324423.

Whitehead, Charles K. 2012. "The Goldilocks Approach: Financial Risk and Staged Regulation." *Cornell Law Review* 97 (5): 1267–1308.