

University of Pennsylvania Carey Law School

## Penn Law: Legal Scholarship Repository

---

Faculty Scholarship at Penn Law

---

5-29-2018

### Regulation and the Marginalist Revolution

Herbert J. Hovenkamp

*University of Pennsylvania Carey Law School*

Follow this and additional works at: [https://scholarship.law.upenn.edu/faculty\\_scholarship](https://scholarship.law.upenn.edu/faculty_scholarship)



Part of the Antitrust and Trade Regulation Commons, Courts Commons, Economic History Commons, Growth and Development Commons, Industrial Organization Commons, Inequality and Stratification Commons, Intellectual Property Law Commons, Law and Economics Commons, Public History Commons, and the Technology and Innovation Commons

---

#### Repository Citation

Hovenkamp, Herbert J., "Regulation and the Marginalist Revolution" (2018). *Faculty Scholarship at Penn Law*. 1984.

[https://scholarship.law.upenn.edu/faculty\\_scholarship/1984](https://scholarship.law.upenn.edu/faculty_scholarship/1984)

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact [PennlawIR@law.upenn.edu](mailto:PennlawIR@law.upenn.edu).

## REGULATION AND THE MARGINALIST REVOLUTION

Herbert Hovenkamp\*

**PRELIMINARY DRAFT: PLEASE CONSULT AUTHOR BEFORE CITING**

### Abstract

The marginalist revolution in economics became the foundation for the modern regulatory State with its “mixed” economy. Marginalism, whose development defines the boundary between classical political economy and neoclassical economics, completely overturned economists’ theory of value. It developed in the late nineteenth century in England, the Continent and the United States. For the classical political economists, value was a function of past averages. One good example is the wage-fund theory, which saw the optimal rate of wages as a function of the firm’s ability to save from previous profits. Another is the theory of corporate finance, which assessed a corporation’s worth by looking at how much capital had been paid in. Marginalism substituted forward looking theories based on expectations about firm and market performance. The optimal rate of wages became the laborer’s expected contribution to the value of the employer; and the value of the corporation became the firm’s anticipated profits. Marginalism swept through university economics in the United States, and by 1920 or so virtually every academic economist was a marginalist.

As a theory of value, marginalism was much more realistic than classical political economy about how market actors behave. At the same time, however, valuations based on expectations about the future, which necessarily included risk, required both more technical analysis and the accommodation of more uncertainty. These changes had a powerful effect on the development of modern regulatory policy in the United States.

Marginalism upended many of the classical conceptions about the market, including assumptions about their robustness, as well as the need for regulation and the optimal type. For regulatory policy the most important issues were: (1) The fixed-cost controversy and the scope of natural monopoly; (2) cost classification, incentives, and ratemaking; (3) the changing domain of market failure (4) market diversity and the rise of sector regulation; (5) deregulation; (6) concerns about the distribution of wealth; and (6) the assessment of risk. The final section examines risk management under marginalism by looking at two diverse but important areas: negligence and products liability in tort law, and administrative review of patents by the Patent Trial and Appeal Board.

---

\*James G. Dinan University Professor, Univ. of Pennsylvania School of Law and The Wharton School. A shorter version of this paper is intended for presentation at a program entitled “New Perspectives on U.S. Regulatory History: Past and Present of Public Utilities and Antitrust Law,” June 4-5, 2018, Harvard Business School. Thanks to Erik Hovenkamp for commenting on portions of a draft.

## Introduction

Of all the effects of the neoclassical revolution in economics, few have been more profound and lasting as its impact on the theory and policy of regulation. The marginalist revolution in economics became the foundation for the modern regulatory State with its “mixed” economy. The so-called “marginal revolution,” which best identifies the dividing line between classical and neoclassical economics, originated in the work of William Stanley Jevons in England, Carl Menger in Austria, Leon Walras in Switzerland, and a little later John Bates Clark in the United States.<sup>1</sup> The initial work was done in the 1860s and 1870s. In the United States the dispute over marginalism led to an upheaval in the American Economic Association, which had been founded in 1885.<sup>2</sup> The neoclassicists eventually won out and marginalist approaches swept the field. Today every mainstream economist is some kind of marginalist.

Formulating marginalism into what became modern price theory and industrial organization became the task of a second generation of marginalists -- most notably Alfred Marshall of Cambridge, whose *Principles of Economics* was published in 1890 and went through numerous editions.<sup>3</sup> Marginalism completely upended classical political economy’s theory of value, which had largely been drawn from averages of past experience. For example, the optimal rate of wages was dictated by the wage-fund doctrine, which held that the fund available for the payment of wages was determined by the amount of capital retained from previous production. This then had to be divided among the number of workers.<sup>4</sup> If a firm’s wages exceeded this amount it would end up borrowing against the future, which was a road to insolvency. An analytically similar classical doctrine was the theory of corporate finance that the value of a corporation equaled the amount of capital that had been paid into the corporation.<sup>5</sup> The marginalist critique of these doctrines observed that they had little to do with actual market behavior. The wage that an employer is willing to pay is not based on the amount in some historical fund, but on the employer’s expectations about how much value the employee will contribute to the firm.<sup>6</sup> Equally, the value of a corporation is not a function of the amount of capital that has been paid in, but rather the firm’s prospects for the future. As one early twentieth century treatise on corporate finance observed concerning the change, “there is very seldom even a close correspondence between the original investment of capital and the value of a

---

<sup>1</sup>See HERBERT HOVENKAMP, *THE OPENING OF AMERICAN LAW: NEOCLASSICAL THOUGHT, 1870-1970* at 27-32 (2015); MARK BLAUG, *ECONOMIC THEORY IN RETROSPECT* 277-310 (3d ed. 1997). See WILLIAM STANLEY JEVONS, *THEORY OF POLITICAL ECONOMY* (1871); CARL MENGER, *PRINCIPLES OF ECONOMICS* (1871); LEON WALRAS, *ELEMENTS OF PURE ECONOMICS* (1874); JOHN BATES CLARK, *THE DISTRIBUTION OF WEALTH: A THEORY OF WAGES, INTERESTS AND PROFITS* (1899).

<sup>2</sup>See HOVENKAMP, *OPENING*, *supra* note \_\_ at 75-76.

<sup>3</sup>ALFRED MARSHALL, *PRINCIPLES OF ECONOMICS* (1890).

<sup>4</sup>HERBERT HOVENKAMP, *ENTERPRISE AND AMERICAN LAW, 1836-1937* at 193-198 (1991).

<sup>5</sup>HOVENKAMP, *OPENING*, *supra* note \_\_, 159-171.

<sup>6</sup>*E.g.*, CLARK, *DISTRIBUTION OF WEALTH*, *supra* note \_\_ (developing marginal productivity theory of wages).

corporation's assets."<sup>7</sup> These changes in economics' methodology of evaluation reflected the general rule of marginalism that value is based on a maximizing actor's willingness to pay or to accept for a given good or opportunity. While the classical theory of value looked backward, the marginalist theory looked forward.

Marginalism upended many of the classical conceptions about the market and, in the process, about the need for regulation and the optimal type. For regulatory policy the most important issues were: (1) The fixed-cost controversy and the scope of natural monopoly; (2) cost classification, incentives, and ratemaking; (3) the changing domain of market failure (4) market diversity and the rise of sector regulation; (5) deregulation; (6) concerns about the distribution of wealth; and (6) the assessment of risk.

The marginalist revolution in economics made modern, sector-specific agency regulation inevitable, at least within the framework that marginalist economics chose.<sup>8</sup> To be sure, the development of regulatory agencies is historically contingent. One can also imagine a world in which we have fewer of them, or their jurisdictional boundaries over subject matter or geography differ from the arrangements that we have. The fact is, however, that the complexity of issues created by marginalism, including accounting for risk and distributional concerns, made them essential nonetheless. It is one thing to roll back the clock on regulation; it is quite another to roll it back on marginalism itself.

One of marginalism's most threatening aspects was that it imposed a significant measure of subjective choice into policy making. The conception of markets in nineteenth century political economy was that they were an objective part of the laws of nature. Classical political economists sometimes spoke of the laws of the economy as a part of natural law.<sup>9</sup> A popular nineteenth century American text initially written by Brown University's Francis Wayland and later co-authored with Aaron Chapin declared that:

Political Economy is that branch of Social Science which treats the production and application of wealth to the well-being of men in society. It is a branch of true science.... By Science, as the word is here used, we mean a Systematic arrangement of the laws which God has established, so far as they have been discovered, of any department of human knowledge.<sup>10</sup>

Marginalism and the resulting regulation were stunningly different from this conception of economic science. Nearly every decision to regulate leads to a menu of policy choices. Different jurisdictions make different choices, and it is not always clear that one works better than another. Further, different approaches may work better in different environments. Regulation often invites experimentation and necessarily criticism about the amount or the particular direction regulation has taken. This rhetoric continues to appear in the debate between

---

<sup>7</sup>WILLIAM HENRY LOUGH AND FREDERICK WILLIAM FIELD, CORPORATION FINANCE: AN EXPOSITION OF THE PRINCIPLES AND METHODS GOVERNING THE PROMOTION, ORGANIZATION AND MANAGEMENT OF MODERN CORPORATIONS 125, 127, 128, 130-131 (1909; 7th ed. 1916).

<sup>8</sup> See discussion *infra*, text at notes \_\_\_\_.

<sup>9</sup>*E.g.*, HENRY WOOD, THE POLITICAL ECONOMY OF NATURAL LAW 18-19 (1894).

<sup>10</sup>FRANCIS WAYLAND, THE ELEMENTS OF POLITICAL ECONOMY 348 (recast by Aaron L. Chapin, 1886).

the advocates for regulation and its detractors. Just to give one example, the perceived need for regulation of the business corporation and its shareholders or creditors began to loom large during the Gilded Age, particularly in the “watered stock” scandals. The regulatory experiments that resulted included state “blue sky” laws,<sup>11</sup> and in the federal government the U.S. Industrial Commission (1898-1902), the Federal Bureau of Corporations (1903-1915), and eventually the Securities Exchange Act of 1934.

Also inherent in marginalism’s emphasis on expected value was greatly increased use of mathematics. As a result, the initial division between traditionalists and marginalists was generational. Older economists lacked the training in mathematics that the new economics required, and they naturally resisted it.<sup>12</sup> The models that marginalist economists developed took expected value and risk into account. Because pricing decisions are made at the margin, neoclassical economics provided a technical basis for dividing the behavior of markets or firms into shorter or longer “runs” or time periods. It produced technical cost classifications into long-run and short-run costs, including the development of the marginal cost curve, perhaps the most distinctive mathematical feature of marginalism. The marginal revenue curve came a generation later.<sup>13</sup>

Closely related was increased attention to the theory of competition, and increasing awareness that “perfect” competition is the exception rather than the rule. Rather, competition exists in degrees, depending mainly on the number of firms and the types of costs that each firm faces.<sup>14</sup> The full implications had to await the early 1930s and the work of Robinson and Chamberlin on imperfect and monopolistic competition.<sup>15</sup> Joseph Schumpeter had the prescience to see this in the mid-thirties, in a review of Joan Robinson’s *Economics of Imperfect Competition*.<sup>16</sup> Prior to that time economists had treated competition and monopoly as the only market structures worth systematic examination, and everything in between as “an academic curiosity.” However, if this intermediate range should emerge as dominant then the traditional laissez faire assumption that markets should be left alone could cease to hold true. Instead, the

---

<sup>11</sup>See, e.g., William W. Cook, “Watered Stock” -- Commissions -- “Blue Sky Laws” -- Stock Without Par Value, 19 MICH. L. REV. 583, 584 (1921) (prominent Gilded Age corporate law treatise author advocating state blue sky statutes). See also LOUIS LOSS & EDWARD M. COWETT, BLUE SKY LAWS (1958); Paul G. Mahoney, *The Origins of the Blue-Sky Laws: A Test of Competing Hypotheses*, 46 J. L. ECON. 229, 230-232 (2003).

<sup>12</sup>HOVENKAMP, OPENING, *supra* note \_\_\_ at 75-76.

<sup>13</sup>The construction of the marginal revenue curve is generally attributed to Joan Robinson, occasionally with Edward H. Chamberlin as an independent discoverer. See Alfred S. Eichner, *Joan Robinson’s Legacy*, 27 CHALLENGE 42 (1984). See JOAN ROBINSON, *THE ECONOMICS OF IMPERFECT COMPETITION* 182-196 (1933; 2d ed. 1969); EDWARD H. CHAMBERLIN, *THE THEORY OF MONOPOLISTIC COMPETITION* 178-193 (1933).

<sup>14</sup>George J. Stigler, *Perfect Competition, Historically Contemplated*, 65 J. Pol. Econ. 1 (1957).

<sup>15</sup>ROBINSON, *supra*; CHAMBERLIN, *supra*.

<sup>16</sup>Joseph A. Schumpeter & A.J. Nichol, *Robinson’s Economics of Imperfect Competition*, 42 J. POL. ECON. 249, 250-251 (1934).

circumstances under which governmental action could increase welfare “becomes so extended as to make these cases the rule rather than more or less curious exceptions.”<sup>17</sup>

Inextricably related was the concept of uncertainty – and relatedly, risk – which developed naturally out of theories that identified value with future expectations rather than past averages.<sup>18</sup> A good example is the revolution in corporate finance.<sup>19</sup> Under the classical theory any judge who knew a little arithmetic and basic concepts of property valuation could decide whether stock was “watered,” which meant that its stated par value exceeded the actual value historically paid in.<sup>20</sup> Answering that question required a judge to determine the amount of paid in cash and noncash capital and divide by the number of shares. The biggest problem was overstatement of the value of noncash property.<sup>21</sup> But the marginalist theory that a firm’s value is a function of anticipated profitability was far more complex, requiring considerable information about the firm itself and its products and management, the market in which the firm operated, and the valuation of uncertain future events.

The initial effect of marginalism was to weaken the classical view that nearly all markets work well, and opinions about the robustness of markets differed a great deal depending on assumptions about such things as the distribution of fixed and variable costs, the flow of information, or the degree and nature of risk or uncertainty that the firm faced. This in turn led increasingly to the view that markets differ from one another, and that many would work better if the state applied a corrective. In other words, marginalism provided a rationale for both vastly increased amount of government intervention in the economy and the idea that the type of intervention should vary from one market to another. As with many revisionist ideas, its early history reflected positions that were later thought to be extreme once marginalism became more normalized in economic and regulatory theory.

### **Fixed Costs, Equilibrium, and Natural Monopoly**

One of the most important issues confronting neoclassical regulatory theory was determining the range of industries and firms to which price regulation should be applied. Policy makers needed to know whether an industry equilibrium with satisfactory results could be achieved without state intervention.

Among Alfred Marshall’s most significant contributions to neoclassical price theory was his development of partial equilibrium analysis, which examined a portion of the economy limited to a single “commodity” or finite time period. Marshall realized that in an economy everything

---

<sup>17</sup>*Id.* at 256.

<sup>18</sup> The most prophetic was FRANK H. KNIGHT, *RISK, UNCERTAINTY, AND PROFIT* (1921).

<sup>19</sup> See discussion *supra*, text at notes \_\_.

<sup>20</sup> *E.g.*, Appeal of Lehigh Ave. Rwy. Co., 499, 129 Pa. 405 (1889) (stock certificate “stands in the hands of the subscriber for so much as, and no more than, the amount actually paid upon it”).

<sup>21</sup> *E.g.*, Gillett v Chicago Title & Trust Co., 230 Ill 373, 82 N.E. 891, 904-05 (1907) (promoters contribution of an unwritten play and an unpatented invention not worth the \$2 million evaluation placed on them. Cf. Carr v. Le Fevre, 27 Pa. 413 (1856) (contribution of coal mining property at appraised value permitted, given that there was no evidence suggesting that the appraisal was fraudulent).

affects everything else. Nevertheless, he argued forcefully for the importance of studying a specific product, assuming that changes in demand and supply for that product had no effect on other goods. In defending this approach he wrote:

“The forces to be dealt with are, however, so numerous that it is best to analyse a few at a time and to work out a number of partial solutions as auxiliaries to our main study. Thus we begin by isolating the primary relations of supply, demand and price in regard to a particular commodity. We reduce to inaction all other forces by the phrase ‘other things being equal’. We do not suppose that they are inert, but for the time we ignore their activity. This scientific device is a great deal older than science; it is the method by which consciously or unconsciously sensible men dealt from time immemorial with every difficult problem of everyday life.<sup>22</sup>

Famously, however, Marshall was unable to get an equilibrium in a competitive market if the firms in it experienced significant fixed costs.<sup>23</sup> Competition would drive prices to marginal cost, which covers only variable costs. The firms in such a market would be forced to charge prices so low that they could not stay in business.

One result of Marshall’s dilemma was that economists became embroiled in a significant “fixed cost” controversy, in which economists and lawyers debated both the meaning and the policy implications of high fixed costs.<sup>24</sup> In antitrust for example, firms argued that in industries with high fixed costs competition would be “ruinous,” forcing firms out of business until only a monopolist survived. Otherwise they would be forced to collude or merge.<sup>25</sup> The economics literature produced many studies involving long- vs. short-period sales, the availability of price discrimination, or other theories attempting to explain how fixed costs could be consistent with stable competition.<sup>26</sup>

The policy responses were less extreme than the literature. First, the Supreme Court consistently rejected “ruinous competition” as a defense in antitrust cases.<sup>27</sup> The principal example of an industry that came under price regulation as a result of the controversy was the railroads, where fixed costs were so high that they were thought to create natural monopolies across a wide range of situations. The issue was also raised in the cast-iron pipe price fixing case, but both the Sixth Circuit in Judge Taft’s well-known opinion and the Supreme Court

---

<sup>22</sup>ALFRED MARSHALL, *PRINCIPLES OF ECONOMICS*, *preface*, xiii (8<sup>th</sup> ed. 1920).

<sup>23</sup>*See, e.g.*, Renee Prendergast, *Increasing Returns and Competitive Equilibrium – the Content and Development of Marshall’s Theory*, 16 *CAMB. J. ECON.* 238 (1992).

<sup>24</sup>For an account of the debate in the United States and its affect on developing antitrust policy, *see* HERBERT HOVENKAMP, *ENTERPRISE*, *supra* note \_\_ at 308-322.

<sup>25</sup>*See, e.g.*, Eliot Jones, *Is Competition in Industry Ruinous*, 34 *Q.J. ECON.* 473 (1920) (summarizing the theory of the day).

<sup>26</sup>*E.g.*, JOHN M. CLARK, *STUDIES IN THE ECONOMICS OF OVERHEAD COSTS* (1923); Spurgeon Bell, *Fixed Costs and Market Price*, 32 *Q.J. ECON.* 507, 509–22 (1918); Frank H. Knight, *Cost of Production and Price over Long and Short Periods*, 29 *J. POL. ECON.* 304, 306–10 (1921).

<sup>27</sup>*See* *United States v. Trans-Missouri Freight Assn.*, 166 U.S. 290, 329-330 (1897) (rejecting the defense). *Accord* *United States v. Joint-Traffic Assn.*, 171 U.S. 505, 519-523, 547-48, 569 (1898),

rejected it there as well.<sup>28</sup> There is some reason for believing that the Supreme Court's hostility toward cartels in high fixed cost industries either led to or exacerbated the first great American merger movement, in which many firms with high fixed costs were forced to merge because antitrust law left them unable to collude.<sup>29</sup>

Marshall's own solution to the fixed cost problem was widely regarded as unacceptable. Borrowing from biology, he reasoned that firms were like trees in a forest. The "representative firm" went through a finite life cycle, growing in the earlier part, maturing, and eventually withering away.<sup>30</sup> As a result there would not be durable monopoly but rather ongoing cycling of firms.<sup>31</sup> Marshall wrote in his Eighth edition:

...the very conditions of an industry which enable a new firm to attain quickly command over new economies of production, render that firm liable to be supplanted quickly by still younger firms with yet newer methods. Especially where the powerful economies of production on a large scale are associated with the use of new appliances and new methods, a firm which has lost the exceptional energy which enable it to rise is likely ere long quickly to decay; and the full life of a large firm seldom lasts very long.<sup>32</sup>

This passage indicates that Marshall was aware of the role of innovation ("newer methods") in upsetting equilibrium and perhaps even making the search for equilibrium less important. But Marshall himself did not do down this road.<sup>33</sup>

One weakness in his biological theory of the life of a representative firm, Marshall acknowledged, might be joint stock companies, or corporations, which did not necessarily go through this ageing cycle. However, Marshall persisted in the view that such companies were inherently inferior mechanisms of production, largely because of problems relating to the separation of ownership and control.<sup>34</sup>

Most of Marshall's intellectual descendants did not follow him into this biological forest. Rather they looked for solutions to the fixed cost problem in two quite different places. Both were driven more by the mathematics of marginalism rather than by biological analogies. First

---

<sup>28</sup>United States v. Addyston Pipe & Steel Co., 85 F. 271, 291 (6th Cir. 1898), modified and aff'd, 175 U.S. 211 (1899) (rejecting "ruinous competition" defense).

<sup>29</sup>E.g., NAOMI R. LAMOREAUX, *THE GREAT MERGER MOVEMENT IN AMERICAN BUSINESS, 1895-1904* at 27-33 (1998). See also LESTER TELSER, *ECONOMIC THEORY AND THE CORE* 387 (1978); George Bittlingmayer, *Did Antitrust Policy Cause the Great Merger Wave?*, 28 J. L. & ECON. 77 (1985); George Bittlingmayer, *Decreasing Average Cost: a New Look at the Addyston Pipe Case*, 25 J. L. & ECON. 201 (1982).

<sup>30</sup>See Gerald F. Shove, *The Representative Firm and Increasing Returns*, 40 ECON. J. 94 (1930).

<sup>31</sup>For a good discussion, see Neil Hart, *Marshall's Theory of Value: the Role of External Economies*, 20 CAMBRIDGE J. ECON. 353, 360-362 (1996).

<sup>32</sup>ALFRED MARSHALL, *PRINCIPLES OF ECONOMICS* 239 (8<sup>th</sup> ed. 1920).

<sup>33</sup>That largely awaited Schumpeter. See JOSEPH A. SCHUMPETER, *CAPITALISM, SOCIALISM, AND DEMOCRACY* (1942), particularly its Chapter 7 on "The Process of Creative Destruction."

<sup>34</sup>*Id.* at 253 (shareholders generally unable "to exercise an effective and wise control over the general management of the business"). See *id.* at 254-258 for Marshall's highly suspicious account of corporations, and particularly about the rise of the "trust" in the United States.



was the idea that the long run average cost curve might be U-shaped rather than continuously downward sloping. The second was the availability of product differentiation and, eventually, monopolistic competition.

The proposition of U-shaped long run average cost is simply that scale economies do not produce increasing returns indefinitely. Rather at some point the curve bottoms out and either remains flat or begins to rise. In that case the equilibrium number of firms in a market with fixed costs can be more than one, depending on the size of the market. For example, if average fixed costs decline only to an output level of 1000 and market demand at that level is 10,000 units, then this market has room for up to 10 technically efficient firms. This idea developed gradually in writings by Marshall's successor Arthur C. Pigou, and Francis Edgeworth, Piero Sraffa, Jacob Viner, and others.<sup>35</sup> Pigou in particular showed that when the supply price of the industry is greater than the marginal cost of a firm, that firm would expand until these costs were equalized. When a firm was in equilibrium, its marginal costs would equal the industry's supply curve. Firms could be expected to jockey for growth by competing to reduce costs.<sup>36</sup>

Of course, the fact that a U-shaped cost curve is possible does not mean that it explains every situation. If the market is sufficiently small in relation to the availability of scale economies there still might be room for only one firm. The other major development that addressed the equilibrium problem was the idea that firms differ from one another in geographic<sup>37</sup> and product space. The development of the idea of product differentiation, and most particularly of monopolistic competition theory in the 1930s, very largely ended the fixed cost controversy.<sup>38</sup> Under monopolistic competition firms compete primarily not by cutting price but rather by differentiating their product in order to appeal to specific consumer tastes. Edward Chamberlin was able to show that even with easy entry and high fixed costs it was possible to have an equilibrium if products were differentiated.<sup>39</sup>

The theory of monopolistic competition accomplished two things for regulatory policy. First, it set people's minds at ease that at least in the world of manufactured products rather than commodities, ruinous competition and collusion or merger to monopoly were not inevitable.

---

<sup>35</sup>Piero Sraffa, *On the Relationships Between Cost and Quantity Produced* (translated from French "Sulle relazioni fra costo e quantità prodotta"), 11 ANNALI DE ECONOMIA 277 (1925); Piero Sraffa, *The Laws of Returns Under Competitive Conditions*, 36 ECON. J. 535 (1926); ARTHUR C. PIGOU ECONOMICS OF WELFARE (3d ed. 1929); Jacob Viner, *Cost Curves and Supply Curves* (1931), reprinted in READINGS IN PRICE THEORY 23-46 (George J. Stigler and Kenneth E. Boulding, eds. 1953). See also Jan Horst Keppler and Jerome Lallemand, *The Origins of the U-Shaped Average Cost Curve: Understanding the Complexities of the Modern Theory of the Firm*, 38 HIST. POL. ECON. 733 (2006). They identify the earliest graphic depiction in Francis Y. Edgeworth, *Contribution to the Theory of Railway Rates – IV*, 23 ECON. J. 206 (1913).

<sup>36</sup>Arthur C. Pigou, *An Analysis of Supply*, 38 ECON. J. 238 (1928). See Herbert Hovenkamp, *Coase, Institutionalism, and the Origins of Law and Economics*, 86 IND. L.J. 499 (2011).

<sup>37</sup>On differentiation in geographic space, see Harold Hotelling, *Stability in Competition*, 39 ECON. J. 41 (1929).

<sup>38</sup>See BLAUG, *supra* note \_\_, 375-379.

<sup>39</sup>E.g., John R. Hicks, *Survey of Economic Theory: The Theory of Monopoly*, 3 ECONOMETRICA 1 (1935); Arthur Smithies, *Equilibrium in Monopolistic Competition*, 55 Q.J. ECON. 95 (1940).

Rather, this became a problem of degree. Second, however, was the realization that monopolistic competition was not so perfect either. Prices were higher than marginal cost. Further, firms under monopolistic competition dedicated considerable resources toward product differentiation, leading to the view that product differentiation was “excessive.” This led to discussions of such topics as whether annual style changes in the automobile industry should be regarded as an unfair method of competition.<sup>40</sup> Relatedly, the intellectual property laws were thought to be harmful to the extent that they facilitated such differentiation. The decades following the New Deal were characterized by a rapid expansion of antitrust policy accompanied by considerable hostility toward intellectual property rights, most notably patents and trademarks.<sup>41</sup>

In any event, the ruinous competition problem itself became largely relegated to common carriers such as the railroads, public utilities, and other transport firms. It has reappeared from time to time in antitrust price fixing case law, but only to be rejected. The most recent is the Apple *eBooks* decision, which rejected the publishers’ argument that ruinous competition in electronic books justified their collective agreement to raise prices and impose these on Amazon.<sup>42</sup>

One important but overhyped idea that originated in the 1960s and 1970s was that even natural monopoly franchises could operate competitively if competition were seen as being “for” rather than “in” the market. Even if a market has room for only one seller at a time, competition to be that seller might keep prices at the competitive level without regulation. The basic theory had already shown up in various “potential competition” antitrust cases.<sup>43</sup> It was also implicit in the common law of requirements contracts. For example, a store owner might prefer to have one person plow snow on an as needed basis over the course of a season, rather than taking bids for each snowstorm. Courts had approved such contract since the Gilded Age.<sup>44</sup> Even though there is only one snow plower over the year, the price is competitive because prospective sellers must bid and rebid for that contract, or franchise.

Harold Demsetz famously queried, why could not the same principle apply to public utilities?<sup>45</sup> Of course, the snow plower has assets such as a truck that can readily be moved from

---

<sup>40</sup> E.g., FTC, *Report on the Motor Vehicle Industry* (1939). The issues are summarized in Note, *Annual Style Change in the Automobile Industry as an Unfair Method of Competition*, 80 *Yale L.J.* 567 (1971). In its main brief to the Supreme Court in the *Brown Shoe* Merger Case, the government argued that rapid style changes in the shoe industry made it very difficult for smaller shoe manufacturers to compete. Brief for the United States, *Brown Shoe Co. v. United States* (Nov. 15, 1961), 1961 WL 101890.

<sup>41</sup> See HOVENKAMP, OPENING, *supra* note \_\_, at 198-200, 227-228, 288-289.

<sup>42</sup> *United States v. Apple, Inc.*, 791 F.3d 290, 333 (2d Cir. 2015), *cert. denied*, 136 S. Ct. 1376 (2016); see Herbert Hovenkamp, *Antitrust and the Design of Production*, 103 *CORNELL L. REV.* \_\_ (2018) (forthcoming), available at <file:///C:/Users/sherm/Downloads/SSRN-id3013122>. See also Herbert Hovenkamp, *Antitrust and Information Technologies*, 68 *FLA. L. REV.* 419, 438-40 (2016).

<sup>43</sup> See, e.g., *United States v. El Paso Natural Gas Co.*, 376 U.S. 651 (1964) (two firms bidding to wholesale natural gas to Los Angeles should be regarded as competitors for merger purposes, even though only one of them had made any sales there).

<sup>44</sup> See HOVENKAMP, OPENING, *supra* note \_\_ at 126-129.

<sup>45</sup> See, e.g., Harold Demsetz, *Why Regulate Utilities?*, 11 *J.L. & ECON.* 55 (1968).

one site to another, permitting the market for plowing snow on a particular parking lot to be very competitive, provided that the prospective snow plowers behave competitively. As soon as one begins to speak about utilities that have durable and nonmoveable installations, such as power, cable, or gas lines, the problem becomes much more complex. The winner of the first round would have a significant bidding advantage over any prospective entrant who would have to install the system. Alternatively, some mechanism might be established for transferring the system from the incumbent to the new winner, or perhaps a government such as a state or municipality could own the system, with successive firms simply agreeing to be its operators.<sup>46</sup>

William Baumol, one of the most optimistic of contestable market proponents, proclaimed it to be “an uprising of in the Theory of Industry Structure.”<sup>47</sup> To be sure, the theory of contestable markets was an important contribution to regulatory theory. However, as a policy matter it never lived up to expectations. Much of the technical theory turned into a discussion of cost classification – mainly, which costs were variable, which were fixed, and which were “sunk.” The later type was crucial. Even a high fixed cost need not interfere with contestability if the asset in question, such as the snow plowing truck, is readily and costless transferrable from one market to another. For obvious reasons, much of the theory focused on the airline industry: although aircraft are costly and durable, if they can easily be transferred from one market to another in response to diverging prices the result should be competition. But this theory of “hit and run” entry largely overlooked details such as the very substantial cost of airports, gate space leases, and other items that were not so readily transferable.<sup>48</sup> Institutions, it seems, always get in the way. The effect of deregulation of airlines has not been contestability, but rather imperfect competition or oligopoly among competitors, with frequent claims of price fixing. The literature generally shows an inverse correlation between price and the number of carriers on a route – which is inconsistent with contestability but quite consistent with oligopoly or collusion.<sup>49</sup> From a consumer welfare perspective, price competition among competing carriers is very likely much superior to the previous regime of price regulation by the now defunct Civil Aeronautics Board, but it is a far cry from contestability.

---

<sup>46</sup>Oliver E. Williamson, *Franchise Bidding for Natural Monopolies—in general and with Respect to CATV*, 7 BELL J. ECON. 73 (1976).

<sup>47</sup>William J. Baumol, *Contestable Markets: An Uprising in the Theory of Industry Structure*, 72 AM. ECON. REV. 1 (1982). The theory was expanded and later published in WILLIAM J. BAUMOL, JOHN C. PANZAR, & ROBERT D. WILLIG, *CONTESTABLE MARKETS AND THE THEORY OF INDUSTRY STRUCTURE* (1988).

<sup>48</sup>See General Accounting Office, *Airline Competition: Effects of Airline Market Concentration and Barriers to Entry on Airfares* (April 1991), available at <https://www.gao.gov/assets/220/214117.pdf> (under deregulation airline markets still exhibit higher prices as they become more concentrated, indicating lack of contestability).

<sup>49</sup>See, e.g., Federico Ciliberto, Eddie Watkins, and Jonathan W. Williams, *Two Screening Tests for Tacit Collusion: Evidence from the Airline Industry* (July 29, 2017), available at <https://www.gao.gov/assets/220/214117.pdf>; [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3012580](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012580); Gaurab Aryal, Federico Ciliberto, & Benjamin T. Leyden, *Public Communication in the Airline Industry* (Feb. 2018, Center for Economic Policy Research, DP12730).

### Cost Classification, Incentives, and Return-based Ratemaking.

Under marginalism, the principal rationale for rate regulation moved from a focus on “public service” to one that focused on the nature of costs. The historical rationale for rate regulation had been that certain firms were “affected with the publick interest,” in the words of Justice Matthew Hale, because of their strategic location. He made that argument in reference to strategically located seaports which held effective monopolies in their service area.<sup>50</sup> The United States Supreme Court adopted it to uphold state price regulation of grain elevators strategically located along railroad tracks.<sup>51</sup> Blackstone’s justification for regulation used the term “prerogatives of the Crown” as opposed to “common callings.”<sup>52</sup> Anyone could engage in a common calling, but only the Crown could authorize someone to exercise one of its prerogatives, which became the subject of exclusive grants.<sup>53</sup>

What distinguished these special firms was both that they were essential gateways to commerce and that competition was not believed to be possible because the market terrain, whether physical or economic, permitted only one firm. Under marginalism, this rationale shifted to the one developed in the previous section – namely, that the firms were natural monopolies, typically with very high fixed costs. As a result, the market itself would produce either ruinous competition or monopoly. More technically, the long run average cost curved sloped downward continuously for such a distance that the optimal number of sustainable firms in a market was one. That fact provided a rationale for price regulation, leaving the question of how to select the proper price.

This neoclassical definition was significantly narrower than the common law classification scheme, which had historically recognized regulated monopoly status for things that were almost certainly not natural monopolies, but merely sellers that were regarded as deserving protection from “excessive” competition. For example, according to Blackstone, at common law a retail market authorized by the local Lord was entitled to protection from

---

<sup>50</sup>Sir Matthew Hale, *De Portibus Maris*, in A TREATISE IN THREE PARTS (c. 1670), reprinted in 1 A COLLECTION OF TRACTS RELATIVE TO THE LAW OF ENGLAND 45 (F. Hargrave ed. 1787).

<sup>51</sup>*Munn v. Illinois*, 94 U.S. 113, 126 (1876), quoting *de Portibus Maris*:

A man, for his own private advantage, may, in a port or town, set up a wharf or crane, and may take what rates he and his customers can agree for cranage, wharfage, housellage, pesage; for he doth no more than is lawful for any man to do, viz., makes the most of his own. . . . If the king or subject have a public wharf, unto which all persons that come to that port must come and unlade or lade their goods as for the purpose, because they are the wharfs only licensed by the queen, . . . or because there is no other wharf in that port, as it may fall out where a port is newly erected; in that case there cannot be taken arbitrary and excessive duties for cranage, wharfage, pesage, &c., neither can they be enhanced to an immoderate rate; but the duties must be reasonable and moderate, though settled by the king's license or charter. For now the wharf and crane and other conveniences are affected with a public interest, and they cease to be *juris privati* only. . . .

<sup>52</sup>1 WILLIAM BLACKSTONE COMMENTARIES ON THE LAWS OF ENGLAND 78 (London 1765-1769).

<sup>53</sup>See William Letwin, *The English Common Law Concerning Monopolies*, 21 UNIV. CHI. L. REV. 355 (1954); Barry E Hawk, *English Competition Law Before 1900*, \_\_ ANTITRUST BULL. \_\_ (2018) (forthcoming), available at <file:///C:/Users/hhovenka/Downloads/SSRN-id3116867>.

competition for a seven-mile radius from its location, which was one third of a day's journey.<sup>54</sup> As late as 1827 the Court of King's Bench held it unlawful for the defendant to sell fish from his house within the specified distance from an authorized local market.<sup>55</sup> New York Chancellor Kent opined in an 1812 decision involving a state issued steamboat patent that the government prerogative to grant exclusive charters for "beneficial public purposes" was clear and that "all our bank charters, turnpike, canal and bridge companies, ferries, markets, etc, are grants of exclusive privileges...." He also acknowledged that the extent of these grants might be "inexpedient or unwise," but that did not undermine the state's power to grant them.<sup>56</sup> His list included some things that could have been natural monopolies (canals and bridges) but others (banks and markets) that almost certainly were not.

Exclusive grants to public service companies in the United States frequently came along with price regulation. Prior to the Gilded Age, however, the prices were typically stipulated in the grantee's charter rather than applied by legislation.<sup>57</sup> In its 1837 *Charles River Bridge* decision the plaintiff argued that government-mandated spatial separation of toll bridges was necessary in order to prevent ruinous competition between them.<sup>58</sup> The plaintiff's charter also stipulated the tolls that it could charge to those crossing the bridge.<sup>59</sup> These two elements of the modern public utility – a monopoly grant and regulated rates – were thus in place long prior to the marginalist revolution.<sup>60</sup> The Supreme Court approved *statutory* price regulation in principle in *Munn v. Illinois*, although it did not consider whether the Constitution required any limitations.<sup>61</sup> In *Munn* there was no charter in which rates could be specified, because *Munn &*

---

<sup>54</sup>BLACKSTONE, COMMENTARIES, *supra* note \_\_ at 219.

<sup>55</sup>Mosley v. Walker, 108 E.R. 640 (K.B. 640).

<sup>56</sup>Livingston v. Van Ingen, 9 Johns 507, 576 (Court for the Correction of Errors of New York, 1812).

<sup>57</sup>HOVENKAMP, ENTERPRISE, *supra* note \_\_ at 105-130.

<sup>58</sup>Proprietors of the Charles River Bridge v. Proprietors of the Warren Bridge, 36 U.S. (11 Pet.) 430, 442 (1837).

<sup>59</sup>STANLEY KUTLER, PRIVILEGE AND CREATIVE DESTRUCTION: THE CHARLES RIVER BRIDGE CASE 10, 32-45 (1971). The charter stipulated:

Each foot passenger (or one person passing), two-thirds of a penny; one person and horse, two pence two-thirds of a penny; single horse cart or sled, or sley, four pence; wheelbarrows, hand-carts, and other vehicles capable of carrying like weight, one penny, one-third of a penny; single horse and chaise, or sulkey, eight pence; coaches, chariots, phaetons and curricles, one shilling each; all other wheel carriages or sleds drawn by more than one beast, six pence; meat cattle and horses passing the said bridge, exclusive of those rode or in carriages or teams, one penny, one-third of a penny; swine and sheep, four pence for each dozen, and at the same rate for a greater or less number; and in all cases the same toll shall be paid for all carriages and vehicles passing the said bridge, whether the same be loaded or not loaded; and to each team one man and no more shall be allowed as a driver to pass free from payment of toll, and in all cases double toll shall be paid on the Lord's day; and at all times when the toll gatherer shall not attend his duty the gate or gates shall be left open.

<sup>60</sup>On the legal evolution *see* HERBERT HOVENKAMP, ENTERPRISE, *supra* note \_\_ at 125-130.

<sup>61</sup>*Munn v. Illinois*, 94 U.S. 113 (1877).

Scott was a common law partnership.<sup>62</sup> In the *Spring Valley Water* case in 1884 the Supreme Court again approved statutory rate regulation, in this case of water rates by a privately operated water works.<sup>63</sup> The statute required that water rates be “reasonable,” and created a commission to determine reasonableness.<sup>64</sup> In approving this procedure, however, the court stated that it was reserving judgment on the possibility that such a commission might set a ‘manifestly unreasonable’ rate.<sup>65</sup> Finally, in 1898 the Supreme Court held that the Fourteenth Amendment prohibited a state from imposing a railroad rate that generated less than “fair value” on the railroad’s property.<sup>66</sup>

While the Interstate Commerce Commission had been created in 1887, it did not have full authority to set rates, preempting state law, until 1920.<sup>67</sup> Pending that, most rate regulation befell the states. The statutory or administrative setting of rates in price regulated industries had always involved some kind of notion that rates must be “above cost,” sufficient to give the regulated firm a reasonable rate of return.

What marginalist price theory provided was a cost classification system for establishing the Constitution’s minimum standard of reasonableness. Aside from sporadic attempts to regulate commodity prices, price regulation prior to the Civil War was largely ad hoc, negotiated on an individual firm basis. By contrast, neoclassical price theory sorted industries by cost structure. In 1887, the same year that the Interstate Commerce Act was passed, Henry Carter Adams wrote an important article classifying industries into three groups: those with decreasing returns to scale, those with constant returns to scale, and those with increasing returns to scale. For industries in the third classification, he argued, competition could not be relied on to make firms perform well. Firms with higher output would have lower costs than smaller ones, leading to a monopoly unless they were constrained. Adams’ principal example was the railroads.<sup>68</sup>

Much of the early neoclassical theory of cost-based price regulation developed in debates over the proper treatment of railroads, an industry that was tailor-made for study of the marginalist economics of costs. The railroads had very high fixed costs for land, track networks, and equipment, but also significant variable, or operating costs, including fuel and labor. The market appeared not to do a very good job of providing the conditions for competitive equilibrium. If a line was too isolated its owners would earn monopoly returns. State officials often responded by chartering additional railroads, resulting in significant overbuilding and

---

<sup>62</sup>See *Wabash, St. L. & Pac. Ry. v. Illinois*, 118 U.S. 557, 569 (1886) (“the case of *Munn v. Illinois* was selected by the court as the most appropriate one in which to give its opinions on that subject; because that case presented the question of a private citizen, or unincorporated partnership”).

<sup>63</sup>*Spring Valley Water-Works v. Schottler*, 110 U.S. 347 (1884).

<sup>64</sup>*Id.* at 353.

<sup>65</sup>*Id.* at 354.

<sup>66</sup>*Smyth v. Ames*, 171 U.S. 361 (1898). See Robert L. Hale, *Does the Ghost of Smyth v. Ames Still Walk?*, 55 HARV. L. REV. 1116 (1942). For a contemporary defense of agency rate regulation, see Adelbert Moot, *Railway Rate Regulation*, 19 Harv. L. Rev. 487 (1906).

<sup>67</sup>Pub. L. No. 66-152, 41 Stat. 456 (1920), codified in scattered sections of 49 U.S.C. (Transportation Act of 1920).

<sup>68</sup>Henry C. Adams, *Relation of the State to Industrial Action*, 1 PUB. AM. ECON. ASSN. 7 (1887).

complaints about “ruinous competition” already in the 1890s.<sup>69</sup> Because most shipping services were fungible, excess capacity tended to drive rates to marginal cost, without enough remaining to amortize fixed costs. In this atmosphere it is no wonder that several economists and legal writers advocated “pooling,” which was essentially cartelization, of competing railroads.<sup>70</sup> Indeed, the Interstate Commerce Commission had approved the very rate setting pool that the Supreme Court condemned in its first antitrust decision on the merits, the *Trans-Missouri* case in 1897.<sup>71</sup>

While collusion would protect the railroads from ruinous competition, it would not protect customers, however. A railroad cartel would charge its profit-maximizing price just as a single firm monopolist would, not the minimum price needed to sustain investment in the industry. In the 1880s Gilded Age scholars of railway regulation such as eventual Yale president Arthur Twining Hadley embarked on a serious economic analysis of railroad costs and pricing.<sup>72</sup> Already by this time, railroad operators and economists understood that paying off fixed cost debt required keeping output high, and that this would be facilitated by permitting price discrimination. Hadley showed how bringing in incremental freight at a lower rate served to reduce overall costs, provided that the rate was greater than incremental operating costs.<sup>73</sup> Hadley was already closing in on a fundamental conception of mid-twentieth century rate regulation:<sup>74</sup> absent capacity constraints the optimal rate brings in every set of customers at the highest rate that class is willing to pay, thus maximizing output. But of course, that was price discrimination which could be trusted to produce loud complaints from those required to pay the higher prices.

The need to price discriminate in order to keep output up led to rate classification rules that seem hysterically complicated and fundamentally unfair to outside observers, although their purpose was clear: any freight rate sufficient to cover the variable costs of shipment would make a positive contribution to net revenue. By bringing in every such customer at the highest price they were willing to pay, the railroad maximized its net revenue and could divide its fixed costs over a larger number of sales. Offsetting this, of course, would be the administrative cost of

---

<sup>69</sup>See discussion *supra*, text at notes \_\_\_.

<sup>70</sup>*E.g.*, Edwin R.A. Seligman, *Railway Tariffs and the Interstate Commerce Law II*, 2 POL. SCI. Q. 369 (1887); Arthur T. Hadley, *The Prohibition of Railway Pools*, 4 Q. ECON. 158 (1890). For further discussion see Herbert Hovenkamp, *Regulatory Conflict in the Gilded Age: Federalism and the Railroad Problem*, 97 YALE L.J 1017 (1988).

<sup>71</sup>*United States v. Trans-Missouri Freight Ass'n*, 58 F. 58, 76-77 (8<sup>th</sup> Cir. 1893), quoting the fourth annual report of the Interstate Commerce Commission, which had approved the arrangement (“To make railroads of the greatest possible service to the country, contract relations would be essential, because there would need to be joint tariffs, joint running arrangements and interchange of cars, and a giving of credit to a large extent, some of which were obviously beyond the reach of compulsory legislation, and, even if they were not, could be best settled, and all the incidents and qualifications fixed, by the voluntary action of the parties in control of the roads respectively.”).

<sup>72</sup>ARTHUR TWINING HADLEY, *RAILROAD TRANSPORTATION: ITS HISTORY AND ITS LAW* (1885).

<sup>73</sup>*Id.* at 117. See also Frank W. Taussig, *A Contribution to the Theory of Railway Rates*, 5 Q. J. ECON. 438 (1891) (showing more technically the efficiency of price discrimination).

<sup>74</sup>See the discussion of Ramsey pricing, *infra*, text at notes \_\_\_.

developing so many classifications, including disputing classification costs with unhappy customers or their rivals.<sup>75</sup> These solutions, extended to other public utilities, produced much of the modern neoclassical economics of industrial organization and price theory, including third-degree price discrimination. For example, the Ramsey pricing solution discussed below contemplated very elaborate rate classifications, literally assessing a rate equal to each individual customer's willingness to pay.<sup>76</sup>

Once marginalism led to classification of costs as fixed, variable, and marginal, the basic theory of rate regulation became simple enough. The devil was in the details. The regulated utility, transportation, or other firm is entitled to a fair rate of return on its fixed cost investment, plus "pass through" of its variable costs. This has come to be called "rate of return" regulation or "cost of service" ratemaking.<sup>77</sup> It has been widely used for a century, but also widely criticized, mainly for providing insufficient incentives for firms to innovate or reduce costs.<sup>78</sup> It also produced a lengthy Constitutional debate about the rate base – namely, whether the firm's return should be based on replacement cost of worn out plant and equipment, or whether it was enough that the firm receive a positive return on its actual historical investment, which was typically much lower.<sup>79</sup> During the Substantive Due Process Era the Supreme Court had insisted on "fair value" and this was generally interpreted by contemporaries as replacement cost.<sup>80</sup> The Court's formulation of the requirements in *Smyth v. Ames* was:

in order to ascertain that [fair] value, the original cost of construction, the amount expended in permanent improvements, the amount and market value of its bonds and stock, the present as compared with the original cost of construction, the probable earning

---

<sup>75</sup> See, e.g., *United States v. Western Pac. R. Co.*, 352 U.S. 59 (1956) (dispute involving higher tariffs for napalm-containing bombs); *Great Northern Ry. Co. v. Merchants' Elevator Co.*, 259 U.S. 285 (1922) (grain).

<sup>76</sup> See discussion *infra*, text at notes \_\_.

<sup>77</sup> The standard treatment is ALFRED E. KAHN, *THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS* (2d ed. 1988). See also JAMES C. BONBRIGHT, ET AL., *PRINCIPLES OF PUBLIC UTILITY RATES* (1988); CHARLES F. PHILLIPS, JR., *THE REGULATION OF PUBLIC UTILITIES: THEORY AND PRACTICE* (1993).

<sup>78</sup> E.g., Harvey Averch & Leland Johnson, *Behavior of the Firm Under Regulatory Constraint*, 52 AM. ECON. REV. 1052 (1962).

<sup>79</sup> The debate became contributed to the early law and economics movement. E.g., Robert L. Hale, *Rate Making and the Revision of the Property Concept*, 22 COLUM. L. REV. 209 (1922) (arguing for historical cost)

<sup>80</sup> E.g., *Smyth v. Ames*, 169 U.S. 466, 541 (1898). The Court indicated that even a rate that provided for full recovery of all costs, including interest and a dividend to shareholders, would be insufficient if less than fair market value:

It cannot, therefore, be admitted that a railroad corporation maintaining a highway under the authority of the state may fix its rates with a view solely to its own interests, and ignore the rights of the public. But the rights of the public would be ignored if rates for the transportation of persons or property on a railroad are exacted without reference to the fair value of the property used for the public, or the fair value of the services rendered, but, in order simply that the corporation may meet operating expenses, pay the interest on its obligations, and declare a dividend to stockholders.

*Id.* at 544.



capacity of the property under particular rates prescribed by statute, and the sum required to meet operating expenses, are all matters for consideration, and are to be given such weight as may be just and right in each case. We do not say that there may not be other matters to be regarded in estimating the value of the property.<sup>81</sup>

Finally in its *Hope Natural Gas* decision (1944), the Supreme Court upheld a rate based on historical cost as the Constitutional minimum.<sup>82</sup>

Two of the most serious criticisms of rate of return regulation are, first, that the basic formulation leads to inefficient pricing to the extent that pass through of variable costs is not a sufficiently close approximation of marginal cost. Second, the administration of cost-of-service ratemaking takes away firms' incentives to innovate and reduce costs.

On the first. Cost of service rate making is not a close approximation to marginal cost pricing. A single price equal to marginal cost would be far too low for most public utilities because it would not compensate the utility for its fixed costs. But how should fixed costs be allocated? The Ramsey solution, which was intended to maximize output and thus reduce the impact of fixed costs, was to price to each customer at the inverse of that customer's elasticity of demand.<sup>83</sup> While Ramsey formalized this solution in the 1920s and it bears his name, railroads had already been doing a rather nontechnical version of this for decades, as Arthur Twining Hadley's 1885 book on railroad rates had elaborated. Regulators understood that the key to profitability was to maximize output on the existing fixed cost network. Any shipper willing to pay more than running expenses was profitable in the short run. At the same time, however, someone had to amortize fixed costs. As a result, the elaborate railroad rate classification schemes made rates for lower value goods such as coal or cement that were far below the rates charged for higher value finished goods. They also charged significantly higher prices per mile for short hauls than for long hauls.<sup>84</sup> The rationale for that phenomenon is readily apparent to anyone who looks at a map of the American railroad network. Often there was only one railroad between two fairly nearby towns. However, as the dropping and pickup points became further apart the network provided more alternatives, making it possible for more railroads to compete for the same shipment. That is, the amount of competition in the market increased as the distance between the two shipping points increased.<sup>85</sup>

---

<sup>81</sup>*Id.* at 546-547.

<sup>82</sup> *FPC v. Hope Natural Gas*, 320 U.S. 591 (1944).

<sup>83</sup> Frank P. Ramsey, *A Contribution to the Theory of taxation*, 37 *ECON. J.* 47 (1927). *See, e.g.*, *Burlington Northern R. co. v. ICC*, 985 F.2d 589, 596 (D.C.Cir. 1993):

Under Ramsey pricing, the regulator allows firms to charge each user a premium over marginal cost in inverse proportion to the elasticity of the user's demand. Because the highest charges fall on the most inelastic demanders, the impact on total usage is minimized. Thus, the Commission believed, it would reconcile the railroad's need for revenue to cover total costs with the least possible distortion of demand (i.e., railroad usage would approximate as nearly as possible the level that would prevail under perfect competition).

<sup>84</sup> *See, e.g.*, ARTHUR T. HADLEY, *RAILROAD TRANSPORTATION: ITS HISTORY AND ITS LAWS* 117 (1885)

<sup>85</sup> *See* Herbert Hovenkamp, *Regulatory Conflict in the Gilded Age: Federalism and the Railroad Problem*, 97 *YALE L.J.* 1017, 1056 (2017).

One idea intended to make regulated pricing resemble market pricing more closely was Ronald Coase's suggestion of two part pricing, which roughly segregated the fixed and variable cost components, permitting marginal cost pricing of the latter.<sup>86</sup> Coase argued that every customer could be required to pay an access fee which was fixed, in the sense that it did not vary with the number of units that the customer used. This fee would be calculated so as to cover the fixed cost components of public utility costs. Then each customer's actual use would be priced out at marginal cost.

Coase's article had actually been written as a response to Harold Hotelling's far more interventionist approach, which was to require the government to pay the fixed cost proportion of transportations and utilities by providing the infrastructure, and then charging customers the marginal cost of operation.<sup>87</sup> Hotelling's argument was highly praised by prominent public utility scholars such as James Bonbright.<sup>88</sup> Coase, who was always fierce in searching out private alternatives, sought one that would minimize government control. Without getting too deeply into the weeds, it is worth noting that this debate, which occurred in the 1930s and 1940s, reveals how central marginalist conceptions had become in economic thinking about regulated pricing, even among those who advocated for a high degree of public ownership.

The second development, which was to have a major impact on the deregulation movement, was Averch's and Johnson's theory of gold-plating.<sup>89</sup> Economically, the "Averch-Johnson effect" operates as a severe qualification on Coase's "The Nature of the Firm" when the firm is in a price-regulated environment.<sup>90</sup> Coase actually conceived of his own article as a first attempt to apply marginalist economics to questions about the optimal size and shape of a business firm. He opened his paper with praise for Marshall:

It is hoped to show ... that a definition of a firm may be obtained which is not only realistic ... but tractable by two of the most powerful instruments of economic analysis developed by Marshall, the idea of the margin and that of substitution, together giving the idea of substitution at the margin.<sup>91</sup>

Coase argued that the firm, driven entirely by the need to maximize its profits, relentlessly compares the marginal cost of doing something internally against the marginal cost

---

<sup>86</sup>Ronald H. Coase, *The Marginal Cost Controversy*, 13 *ECONOMICA* (n.s.) 169 (1946). For elaboration, see 1 KAHN, *ECONOMICS OF REGULATION*, *supra* note \_\_ at 95-100.

<sup>87</sup>Harold Hotelling, *The General Welfare in Relation to Problems of Taxation of Railway and Utility Rates*, 6 *ECONOMETRICA* 242 (1938).

<sup>88</sup>James C. Bonbright, *Major Controversies as the Criteria of Reasonable Public Utility Rates*, 30 *AM. ECON. REV. PAPERS & PROCEEDINGS* 379 (1941) (following Hotelling). See also Nanch Ruggles, *Recent Developments in the Theory of Marginal Cost Pricing*, 17 *REV. ECON. STUD.* 107 (1949).

<sup>89</sup>Harvey Averch & Leland L. Johnson, *Behavior of the Firm Under Regulatory Constraint*, 52 *AM. ECON. REV.* 1052 (1962).

<sup>90</sup>Ronald H. Coase, *The Nature of the Firm*, 4 *ECONOMICA* (n.s.) 386 (1937).

<sup>91</sup>*Id.* at 386.

of procurement from outside, choosing whichever produces the largest payoff. As a result, the firm's structure is "efficient" to the extent that the markets in which it purchases are efficient.<sup>92</sup>

But suppose the firm is guaranteed a profitable price, equal to or slightly above marginal cost, on all of its durable procurements. The firm will have an incentive to perform activities internally rather than through purchase on the market, even though the latter is the better choice in a competitive market. The firm would have an incentive to integrate vertically or to expand into other markets that would not exist absent regulation. To the extent that it gets its rate of return on its capital investment and is limited to pass through of costs for variable cost items such as labor, it has an incentive to invest relatively more in capital assets. As a result, regulated firms tend to be excessively capital intensive when compared with the unregulated market.<sup>93</sup> Perhaps more ominously, the regulated firm would have an incentive to build unneeded infrastructure, knowing that it would be guaranteed a profitable rate of return.<sup>94</sup> By contrast, a competitive firm would enlarge its plant only if anticipated receipts, determined entirely by the market, exceeded anticipated costs.

The Averch-Johnson literature produced interesting collateral issues, such as the regulated utility's right to recover for its "stranded" costs, which are costs incurred in enlargement of infrastructure that later turned out to be unnecessary or unwise.<sup>95</sup> In some cases this happened because of policy changes that made a previous investment, such as a coal-fired plant or hospital's certificate of need,<sup>96</sup> improvident. Importantly, however, firms acting under regulation often propose capital investments to a regulator, and Averch-Johnson suggests this can happen even if those investments would have been inefficient in a competitive market. Then later, when the regulator comes to its senses or the increased demand that the regulated utility promised do not materialize, the regulator withdraws its authorization. This might leave a partially built or recently on line plant "stranded," in the sense that it is no longer needed. The firm may then claim that the regulatory approval and its subsequent retraction is an improper taking of its property without compensation.<sup>97</sup>

---

<sup>92</sup> Note that monopoly in a supply market might induce a firm to do something internally even though external procurement might be cheaper if that market were competitive.

<sup>93</sup>For a simple economic analysis, together with a summary of the most important critiques, see W. KIP VISCUSI, JOHN M. VERNON, & JOSEPH E. HARRINGTON, JR., *ECONOMICS OF REGULATION AND ANTITRUST* 433-436 (4<sup>th</sup> ed. 2005); JEAN-JACQUES LAFFONT & JEAN TIROLE, *A THEORY OF INCENTIVES IN PROCUREMENT AND REGULATION* 91-93 (1993). For more technical treatment see Paul L. Joskow, *Regulation of Natural Monopolies* 1229, in 2 *HANDBOOK OF LAW AND ECONOMICS* (A. Mitchell Polinsky & Steven Shavell, eds., 2007). For deep skepticism, see Stephen M. Law, *Assessing the Averch-Johnson-Wellisz Effect for Regulated Utilities*, 6 *INT'L J. ECON. & FIN.* 41 (2014).

<sup>94</sup>See 1 ALFRED E. KAHN, *THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS* 35-36 (1970).

<sup>95</sup>See Herbert Hovenkamp, *The Takings Clause and Improvident Regulatory Bargains*, 108 *YALE L.J.* 801 (1999).

<sup>96</sup> On the latter, see Michael E. Granfield, *Resource Allocation Within Hospitals: an Unambiguous Analytical Test of the A-J Hypothesis*, 7 *J. APPLIED ECON.* 241 (1975)

<sup>97</sup>See GREGORY SIDAK AND DANIEL F. SPULBUR, *DEREGULATORY TAKINGS AND THE REGULATORY CONTRACT: THE COMPETITIVE TRANSFORMATION OF NETWORK INDUSTRIES IN THE UNITED STATES*

No matter how one feels about Averch-Johnson, one thing that emerges clearly is that cost-of-service rate making is not the best way to determine price and output in any market where competition is viable. As a result, price regulated markets must be pared down so that price and output actually be regulated only in those portions of the market where competition seems not to work. This has been one of the principal consequences of the deregulation movement discussed below.<sup>98</sup>

### **The Changing Domain of Market Failure: Regulation and Public Choice**

The marginalist revolution gave the United States the “mixed” economy that it has today.<sup>99</sup> Under marginalism, markets are not as robust as pre-marginalists believed, but they remain the dominant means by which resources move around. The most important economic policy function of government today is identification of markets that work well when left relatively untended and those that do not, and then coming up with the correct prescription for those in the latter group.

This makes the concept of “market failure” central to modern public decision making. In the political economy of regulation, the term has more than a single meaning. The dominant definition within neoclassical economics relates to the inability of a market to reach an equilibrium on its own, or otherwise to reach only equilibria that exhibit unsatisfactory output and prices.<sup>100</sup> As an example of the latter, when entry is impossible monopoly may be a stable equilibrium, but one in which price and output are suboptimal.

Under the strictest definition, static market failure is any durable deviation from marginal cost pricing. Of course, dynamic factors such as innovation may simultaneously increase the innovator’s price-cost margins while also increasing output or welfare, at least in the long run.<sup>101</sup> For example, one of the longest-standing responses to market failure is the intellectual property system, which involves government creation of exclusive rights to facilitate limited periods of high price-cost margins as an inducement to innovation. In any event, market failure is often defined in ways that do not use Pareto efficiency or perfect competition as a baseline – for

---

(1997). See also Emily Hammond & Jim Rossi, *Stranded Costs and Grid Decarbonization*, 82 BROOK. L. REV. 645 (2017).

<sup>98</sup>See discussion *infra*, text at notes \_\_\_\_.

<sup>99</sup>For a robust defense, see JACOB S. HACKER AND PAUL PIERSON, *AMERICAN AMNESIA: HOW THE WAR ON GOVERNMENT LED US TO FORGET WHAT MADE AMERICA PROSPER* (2017).

<sup>100</sup>The standard discussion is Francis M. Bator, *The Anatomy of Market Failure*, 72 Q. J. ECON. 351 (1958). An effort at a strict definition is Paul A. Samuelson, *The Pure Theory of Public Expenditure*, 36 REV. ECON. & STAT. 387 (1954). For good commentary, see PUBLIC GOODS AND MARKET FAILURES: A CRITICAL EXAMINATION (Tyler Cowen, ed. 2017).

<sup>101</sup>See, e.g., Richard R. Nelson, *Thinking About Technology Policy: “Market Failures” versus “Innovation Systems”* (working paper, UCL Inst. For Innovation and Public Purpose (Oct. 2017), available at <https://www.ucl.ac.uk/bartlett/public-purpose/publications/2018/jan/thinking-about-technology-policy-market-failures-versus-innovation-systems>

example, one can speak of it as “the failure of the market to bring about results that are in the best interests of society as a whole.”<sup>102</sup>

The technical neoclassical definition of market failure generally excludes purely distributive factors. Of course, sometimes income inequality can lead to market failure, particularly when inequality makes it more costly for poorer people to move into more productive occupations,<sup>103</sup> or when lack of competitive pressure on certain groups induces them to be less productive.<sup>104</sup> To the extent maldistribution of wealth impairs productivity, and thus output, it can be a market failure even under a strictly neoclassical definition.

The classical political economists had strong faith in markets. To be sure, there were some qualifications, such as Thomas Malthus population argument that the market would invariably force the population to subsistence levels,<sup>105</sup> Ricardo’s concerns with monopoly in land rents,<sup>106</sup> or perhaps most famously John Stuart Mill’s discussion of natural monopoly and the British postal system or lines for gas lighting.<sup>107</sup> But for the most part the classical political economists regarded these as rare exceptions to the general theory of markets. That was even truer of classical political economy in the United States, where an abundance of undeveloped land was widely seen as making the concerns expressed by Ricardo and Malthus relatively unimportant.<sup>108</sup> Increasingly after the Jackson era, American legislatures and courts trusted markets to allocate resources properly. Some states even invalidated monopoly rights for rather clear natural monopoly industries, such as gas lighting.<sup>109</sup>

A common characteristic of the bridges, toll roads, canals, and railroads that received monopoly charters is that they required a significant fixed cost investment and in most cases one installation would be adequate to handle all of the traffic.<sup>110</sup> As a result, they were true natural monopolies. Competition would drive prices to marginal cost without enough remaining to cover fixed costs. Indeed, in the *Charles River Bridge* case the plaintiffs raised the “ruinous

---

<sup>102</sup>Alain Marciano and Steven G. Medema, *Introduction: Market Failure in Context*, 47 HIST. POL. ECON. 1 (2015).

<sup>103</sup>See, e.g., Kim A. Weeden and David B. Grusky, *Inequality and Market Failure*, 38 AMERICAN BEHAVIORAL SCIENTIST 473 (2013).

<sup>104</sup> For example, if it could be shown that a tax and transfer system that levelled income also increased productivity, then this would be a correction of a market failure in the neoclassical sense.

<sup>105</sup>THOMAS ROBERT MALTHUS, *ESSAY ON THE PRINCIPLE OF POPULATION* (1798).

<sup>106</sup>DAVID RICARDO, *THE PRINCIPLES OF POLITICAL ECONOMY AND TAXATION* (1817).

<sup>107</sup>JOHN STUART MILL, *PRINCIPLES OF POLITICAL ECONOMY* 143 (1848).

<sup>108</sup>Herbert Hovenkamp, *The Political Economy of Substantive Due Process*, 40 STAN. L. REV. 379, 421-431 (1988), summarizing how nineteenth century American political economists believed that the vast amounts of undeveloped land in the United States established that the threats offered by Malthus and Ricardo would occur in the distant future, if at all.

<sup>109</sup> E.g., *Norwich Gas Light Co. v. Norwich City Gas Co.*, 25 Conn. 19 (1856) (invalidating monopoly grants for a gas light utility granted by both the state and the local government, and permitting second firm to install competing lines).

<sup>110</sup>See HOVENKAMP, *ENTERPRISE*, 105-169, 199-206.

competition” defense sixty years before it came into use in railroad cases.<sup>111</sup> Tolls decreased by half to two-thirds upon the opening of the competing Warren Bridge.<sup>112</sup> By the time the case reached the Supreme Court the Charles River Bridge had already closed. It would reopen in 1841 when the Commonwealth turned it into a free bridge.<sup>113</sup>

At least for a time, the marginalist revolution very largely brought the classicists’ robust faith in markets to an end – a view undoubtedly aided by World War I and the Depression. The concerns about fixed costs and failure of equilibrium discussed previously<sup>114</sup> were one significant manifestation, but there were others. The dramatic rise of antitrust was driven by the belief that even markets that are not natural monopolies can fail. Indeed, the extent to which antitrust was brought to bear in *nonmonopolized* markets is striking. After the New Deal Antitrust became a vehicle for controlling manufacturer-created distribution systems in competitively structured markets,<sup>115</sup> as well as controlling price differences between competing dealers,<sup>116</sup> and for limiting the effects of product differentiation. Underlying this was a fear of vertical integration that, at least as a matter of economics, can only be described as hysterical.<sup>117</sup>

Increasingly after the mid-twentieth century, regulatory theory embraced a renewed neoclassicism that once again saw markets as robust, although not as robust as they were perceived by the classicists. Coasean thinking strongly emphasized market solutions, even in markets that did not meet the requirements for perfect competition. Indeed, the Coase Theorem itself focused on bargaining in bilateral monopolies.<sup>118</sup> Entitlements such as communications spectrum<sup>119</sup> or the right to pollute became tradeable and the scope of market failure declined to near non-existence. In the Coasean vision even the lighthouse, frequently given as an example of resistance to free market supply, became freed from government control.<sup>120</sup>

Worse yet, the rise of modern public choice theory in the 1960s turned interest group capture rather than market failure into the positive rationale for regulation. Writers like Buchanan and Tullock began with perfect competition as the baseline, and then sought to explain how democratic political voting produced harmful deviations.<sup>121</sup> Mancur Olson’s much more popular but equally devastating *Logic of Collective Action* also began with competitive markets

---

<sup>111</sup>Proprietors of the Charles River Bridge v. Proprietors of the Warren Bridge, 36 U.S. (11 Pet.) 430, 448, 461 (1837). On the railroads and ruinous competition, see discussion *supra*, text at notes \_\_.

<sup>112</sup>Proprietors of the Charles River Bridge v. Proprietors of the Warren Bridge, 24 Mass. (7 Pick.) 344, 529 (1830).

<sup>113</sup>KUTLER, *supra* note \_\_ at 32-45.

<sup>114</sup>See discussion *supra*, text at notes \_\_.

<sup>115</sup>See HOVENKAMP, OPENING, *supra* note \_\_, Chs. 11, 12; and LAURA PHILLIPS SAWYER, AMERICAN FAIR TRADE: PROPRIETARY CAPITALISM, CORPORATISM, AND THE “NEW COMPETITION,” 1890-1940 (2018).

<sup>116</sup>*Ibid.* (discussing the Robinson-Patman Act, which was passed in 1936. 15 U.S.C. §13).

<sup>117</sup>*Id.*, ch. 11

<sup>118</sup>Ronald H. Coase, *The Problem of Social Cost*, 3 J. L. & ECON. 1 (1960).

<sup>119</sup>E.g., Ronald H. Coase, *The Federal Communications Commission*, 2 J. L. & ECON. 1 (1959).

<sup>120</sup>Ronald H. Coase, *The Lighthouse in Economics*, 17 J. L. & ECON. 357 (1974).

<sup>121</sup>JAMES BUCHANAN AND GORDON TULLOCK, THE CALCULUS OF CONSENT: LOGICAL FOUNDATIONS OF CONSTITUTIONAL DEMOCRACY (1965)

as a baseline and then tried to explain regulatory deviations as political oligopolies or cartels that enriched their organizers at the expense of society.<sup>122</sup> Ironically, Olson borrowed his regulatory oligopoly theory straight from Edward Chamberlin, who had been his dissertation director.<sup>123</sup> Under the Stigler-Posner-Pelzman model of regulation, its roots lay in political bargaining among special interests rather than in anything having much to do with neoclassical price theory or industrial organization theory.<sup>124</sup> Interest group capture became the dominant explanation device for regulation.<sup>125</sup>

Ironically, theories of interest group capture actually entered American politico-economic discourse from the left, through writers such as Charles Beard and Gabriel Kolko. Beard argued during the Progressive Era that the Constitution was largely a product of capture by urban owners of personal property over much more populous but widely dispersed agrarian interests.<sup>126</sup> Kolko, a Marxist, argued in the 1960s that Gilded Age railroad regulation really occurred at the behest of the railroads themselves, who wanted regulation in order to shelter themselves from competition.<sup>127</sup> After that, regulatory capture theory moved rightward.<sup>128</sup>

Today we seem to have reached a balance in regulatory theory in which a combination of traditional price and organization economics and public choice theory are called upon as the best positive explainers of regulatory policy. In a democratic society, to see regulation as nothing more than the consequence of market structure or cost characteristics is naïve, but so is seeing it as nothing more than the outcome of fights among competing interest groups. For example, today in the great majority of states groceries, clothing, and automobiles are sold in markets that are more or less competitive, with price set predominantly through voluntary transactions. By contrast, retail electricity and natural gas are sold through regulated monopolies. Few would seriously argue that this is because electric power lobbyists have more clout than those who work in the food and agricultural markets.

To be sure, interest group theory may explain much of the details of regulatory policy. As noted below, the deregulation movement has done a great service by distinguishing those attributes of the economy that are best left to the market from those that require more

---

<sup>122</sup>MANCUR OLSON, JR., *THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS* (1965).

<sup>123</sup>See Hovenkamp, Opening, *supra* note \_\_ at 313-314.

<sup>124</sup>George J. Stigler, *The Theory of Economic Regulation*, 2 *BELL J. ECON. & MGMT. SCI.* 3 (1971); Richard A. Posner, *The Social Costs of Monopoly and Regulation*, 83 *J. POL. ECON.* 807 (1975); Richard A. Posner, *Taxation by Regulation*, 2 *BELL J. ECON. & MGMT. SCI.* 22 (1971); Sam Peltzman, *Toward a More General Theory of Regulation*, 19 *J.L. & ECON.* 211 (1976); Sam Peltzman, *The Economic Theory of Regulation After a Decade of Deregulation*, in *BROOKINGS PAPERS ON ECONOMIC ACTIVITY: MICROECONOMICS 1* (1989); Sam Peltzman, *An Economic Interpretation of the History of Congressional Voting in the Twentieth Century*, 75 *AM ECON. REV.* 656 (1985).

<sup>125</sup>*E.g.*, *THE REGULATED ECONOMY: A HISTORICAL APPROACH TO POLITICAL ECONOMY* (Claudia Goldin & Gary D. Libecap, eds., 1994).

<sup>126</sup>CHARLES A. BEARD, *AN ECONOMIC INTERPRETATION OF THE CONSTITUTION OF THE UNITED STATES* (1913)

<sup>127</sup>GABRIEL KOLKO, *RAILROADS AND REGULATION, 1877-1916* (1965).

<sup>128</sup>See HOVENKAMP, OPENING, *supra* note \_\_ at 308-314.

governmental decision making.<sup>129</sup> In the process, however, the deregulation movement faces abundant criticism from those who think it has gone too far on the one side, and those who think that it has not gone far enough on the other.

### **Market Diversity and Agency Sector regulation**

Market diversity and market failure go hand in hand. Markets are like Tolstoy's families. Happy families are all pretty much alike, but each unhappy family is unhappy in its own way.<sup>130</sup> When markets work well they require relatively little government attention other than the rules of property and contract, which are very largely common to all. But they fail for reasons that are typically quite specific. Even natural monopolies, while exhibiting a common set of characteristics, have unique features and information requirements that operate on the regulator.

An important consequence of the marginalist revolution was the idea that markets differ from one another, particularly when they fail. As a result, they were thought to require different regulatory fixes. This thesis flowed from two things. First the cost classifications that occupied marginalists, including the emergent theory of regulated monopoly, not only made price and output analysis more complex, it also revealed that while firms within the same market and using the same technology often were more-or-less the same, the firms in one market could have cost structures that were quite different from those in another market. Further, the relationship between fixed costs and market size varied enormously. A relatively small railroad operating within a single state might require price regulation, while giant petroleum or steel refiners might be made to operate differently. Large markets tended to require larger regulators – most notably the federal government if they spilled over more than a single state. Small markets were best controlled by smaller regulators – namely, the states and governmental subdivisions. The diversity was not merely geographic, however. It was also production or technology specific. For example, the cost classification problems that might show up in electric power generation differed from those for natural gas production or telecommunications. Agencies with unique and exclusive jurisdictions were in large part a response to this perception of market diversity, replacing a nineteenth century conception in which most “regulating” was carried on by courts of general jurisdiction.

The other problem leading to agency regulation was risk management.<sup>131</sup> Once again, however, the nature and technical features of the relevant risks could vary widely from one technology or product to another. Classical valuations that depended on the past did not take risk into account because relevant risks were already encountered and included in the calculus. But risk became a significant – indeed, in some cases overwhelming – part of valuation and optimization within the marginalist framework. This is particularly true of activities that produce numerous spillovers or other situations where the market itself or the common law cannot force people to internalize the social harm their activities cause.

---

<sup>129</sup>See discussion *infra*, text at notes \_\_\_\_.

<sup>130</sup>See LEO TOLSTOY, ANNA KARENINA I (1877) (Constance Garnett trans., 2000) (“Happy families are all alike; every unhappy family is unhappy in its own way.”). Thanks to Suzanna Sherry.

<sup>131</sup>See discussion *infra*, text at notes \_\_\_\_.



## Deregulation

The deregulation movement forced reexamination of many issues in the debate over neoclassical market theory and the rationales for government intervention. The movement began in government policy during the waning years of the Carter administration and accelerated during the 1980s.<sup>132</sup> The initial impact of deregulation has been substantial. By one measure, industries characterized by “full” regulation of price and entry produced about 17 percent of the GNP in 1977, but about 6.6 percent in 1988.<sup>133</sup> The deregulation movement was significantly a reaction to excesses in the granting of monopoly status and needless rate-of-return price regulation. Then Professor Stephen Breyer’s 1984 book, *Regulation and its Reform*, pointed out how legal entry restrictions and price regulation were imposed in industries such as trucking that were competitively structured or at least capable of being so.<sup>134</sup>

At the theoretical level the deregulation movement was a response to several developments. An important one was changes in technology.<sup>135</sup> For example, thanks to the development of wireless and advanced switching capabilities, the justifications for a single natural monopoly telephone system disappeared. The real instigators were wireless firms such as MCI and Sprint who convinced courts and the FCC to upset AT&T strenuously defended dominance.<sup>136</sup> More theoretically, it became clear that regulation could be a very costly enterprise in relation to any benefits that it produced. One well known early example of this critique is the evolution of James Landis, Chairman of the SEC during the New Deal. In writings a generation apart Landis initially lauded regulation as the savior of the economy,<sup>137</sup> but later lamented that it had not come close to satisfying expectations.<sup>138</sup>

However, most of the low hanging fruit has already been picked. Further deregulation may threaten the environmental, health, or safety by amounts considerably greater than the cost savings themselves, depending on the industry.<sup>139</sup> That is more likely to be the case as a greater amount of regulation is made subject to mandatory cost-benefit analysis, and as the

---

<sup>132</sup> See, e.g., Motor Carrier Act, 94 Stat. 793 (1980), codified in scattered sections of 49 U.S.C. (removing most rate regulation, except for the obligation to file a rate, and most entry restrictions). A later regulation removed most obligations to file rates. 49 U.S.C. §13710. Rail rates were also largely deregulated. 49 U.S.C. §10502.

<sup>133</sup> See W. KIP VISCUSI, JOHN M. VERNON & JOSEPH E. HARRINGTON, JR., *ECONOMICS OF REGULATION AND ANTITRUST* 316 (3d ed. 2000).

<sup>134</sup> STEPHEN BREYER, *REGULATION AND ITS REFORM* (1984).

<sup>135</sup> See 2A PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶241b1 (4<sup>th</sup> ed. 2015).

<sup>136</sup> E.g., *MCI Commc’ns Corp. v. AT&T Co.*, 708 F.2d 1081 (7th Cir. 1981), cert. denied, 464 U.S. 891 (1983) (applying antitrust’s essential facility doctrine to force AT&T to permit wireless carrier to connect into its system); *in re Carterfone* 13 F.C.C.2d 420 (1968).

<sup>137</sup> JAMES M. LANDIS, *THE ADMINISTRATIVE PROCESS* (1938).

<sup>138</sup> JAMES M. LANDIS, *REPORT ON REGULATORY AGENCIES TO THE PRESIDENT-ELECT* (G.P.O. 1960). See HOVENKAMP, *OPENING*, *supra* note \_\_ at 308-309.

<sup>139</sup> See Ed Dolan, *Is Overregulation Rally Holding Back the U.S. Economy*, *HARV. BUS. REV.* (Jan. 8, 2018)

methodologies of CBA have improved. If a regulation was fully and accurately evaluated under CBA, then removing it by implication will cause more harm than good.

Contestability theory, which was widely touted as opening the path toward broad deregulation,<sup>140</sup> has actually had much less to do with implementing deregulation than have more fundamental concerns about the values of incumbent competition and the economic inefficiency of regulation, particularly those expressed in the Averch-Johnson literature.<sup>141</sup> For the most part, when deregulation has moved markets from regulated monopoly to some sort of competition the competition that emerged has been among incumbent firms, not monopoly incumbents who vied with one another for the right to occupy that position. In other words, the Williamson analysis of contestability, emphasizing the difficulties in the presence of costly and durable infrastructure, has been much more influential on actual policy than the Demsetz analysis.<sup>142</sup> The exceptions have tended to occur in situations where the production facility, network or grid is already in place and the market proceeds by competitive bids for single suppliers during a defined time period.

The greatest focus of deregulation has been on situations where competition among actual incumbents is possible and markets or political institutions can be designed in a way that will accommodate them. The earliest targets for deregulation, which included trucking and passenger airlines, certainly fell into that category. More networked industries eventually followed. For example, the dramatic increases in the amount of competition that have occurred in telecommunications since the 1982 antitrust consent decree<sup>143</sup> and the 1996 Telecommunications Act<sup>144</sup> have largely been increases in competition among multiple incumbent firms. Contestability theory has not had a significant role. AT&T's divestiture of Western Electric led to a fiercely competitive market for devices.<sup>145</sup> For a lengthy period local telephony operated as a price-regulated monopoly, while long distance provision was competitively structured.<sup>146</sup> Today even local service is competitive, although the competition is typically between the resident ILEC<sup>147</sup> and one or more cable or other internet service providers that sell VOIP<sup>148</sup> or perhaps a different technology. Looking at the economy as a whole, true contestability, where a single firm is the current seller in a market with significant infrastructure but must re-bid for that right, is a relative rarity.

---

<sup>140</sup>See discussion *supra*, text at notes \_\_.

<sup>141</sup>See discussion *supra*, text at notes \_\_.

<sup>142</sup>See discussion *supra* note \_\_; and *Williamson, supra* note \_\_; *Demsetz, supra* note \_\_.

<sup>143</sup>*United States v. AT&T Co.*, 552 F. Supp. 131 (D.D.C. 1982), *aff'd sub nom. Maryland v. United States*, 460 U.S. 1001 (1983).

<sup>144</sup>Telecommunications Act of 1996, Pub. L. 104-104, 110 Stat. 56 (1996), codified mainly in scattered sections of 47 U.S.C.

<sup>145</sup>See *United States v. Western Elec. Co., Inc.*, 604 F. Supp. 256, 261 n.23 (D.D.C. 1984), cert. denied, 480 U.S. 922 (1987),

<sup>146</sup>See *AT&T Corp. v. Iowa Utilities Board*, 525 U.S. 366 (1999).

<sup>147</sup>Incumbent Local Exchange Carrier, or the firm designated as the incumbent carrier in a given region.

<sup>148</sup>Voice Over Internet Protocol, which permits real time voice transmission on the internet.

The other thing that has accompanied deregulation is a counterbalancing increase in the use of antitrust law.<sup>149</sup> For example, under comprehensive regulation industries such as the airlines were once regarded as virtually immune from antitrust scrutiny,<sup>150</sup> but hardly today. Substantially, this is the result of a paradigm shift in the theory of regulation, away from the control of markets and toward the control of practices. In most cases of deregulation, regulators do not simply exit from a market. Rather, they apply regulatory tools more selectively. During the 1960s we tended to think of federal agency regulation as “pervasive,” in the sense that it left little discretion over pricing and exclusionary practices to the individual firm, and when these occurred they were usually placed within the jurisdiction of the regulatory agency.<sup>151</sup> Today, however, we view regulation as more “transactional,” in that it applies to a particular practice rather than an entire industry. When a specific transaction is within the jurisdiction of the agency and the agency is actually overseeing it, then antitrust has no role to play. But antitrust can be brought to bear in those areas that the regulators do not tend.<sup>152</sup> This hybrid approach is particularly descriptive of telecommunications, the airlines, healthcare, and to a lesser extent energy.

### Wealth Distribution

The issues swirling around regulation and wealth distribution tend to focus on two concerns. The first is whether government control should be used to redistribute wealth for its own sake, simply because we regard certain distributions as unfair or we think social welfare is greater by some measure as wealth is more evenly distributed. The other is whether maldistribution of wealth affects productivity, and thus redistribution may spur economic growth. The latter one falls within modern neoclassical concerns about regulation. Under most conceptions of economic welfare, the former one does not.

For the first two generations of marginalists, concerns about the distribution of wealth in the economy were paramount. Further, since most did not expect that maldistribution of wealth would right itself through ordinary market forces, some kind of government intervention was thought necessary.

The early marginalists believed in declining marginal utility, but also in the policy value of interpersonal utility comparisons. All things equal, a more even distribution of wealth produced greater welfare. Individuals were thought to experience declining marginal utility for any good, including wealth, as they have more. The leap the early marginalists were willing to

---

<sup>149</sup>See AREEDA & HOVENKAMP, *supra* note \_\_, ¶241

<sup>150</sup>E.g., *Hughes Tool Co. v. TWA, Inc.*, 409 U.S. 363 (1973); *Pan Am World Airways, Inc. v. United States*, 371 U.S. 296 (1963).

<sup>151</sup>See AREEDA & HOVENKAMP, *supra* note \_\_, ¶¶240c1.

<sup>152</sup>See, e.g. *Nat'l Gerimedical Hosp. v. Blue Cross*, 452 U.S. 378 (1981):

To be sure, where Congress did intend to repeal the antitrust laws, that intent governs, but this intent must be clear. Even when an industry is regulated substantially, this does not necessarily evidence an intent to repeal the antitrust laws with respect to every action taken within the industry....Intent to repeal the antitrust laws is much clearer when a regulatory agency has been empowered to authorize or require the type of conduct under antitrust challenge.

make was that this attribute applied across persons as well as to a single individual. For example, Marshall wrote in the third edition of *Principles* that “[A] pound's worth of satisfaction to an ordinary poor man is a much greater thing than a pound's worth of satisfaction to an ordinary rich man.”<sup>153</sup> Marshall’s successor at Cambridge, Arthur C. Pigou, followed: a “transference of income from a relatively rich man to a relatively poor man of similar temperament, since it enables more intense wants to be satisfied at the expense of less intense wants, must increase the aggregate sum of satisfaction.”<sup>154</sup> Many early American marginalists agreed. For example, Chicago economist Jacob Viner acknowledged in 1925 that “Changes in the relative distribution of income as between different classes will bring about changes in the amount of welfare, even though the aggregate real income of the community remains the same.”<sup>155</sup> Within this framework a social welfare scheme that transferred wealth away from the wealthy and toward the poor increased aggregate welfare to the extent that the wealthy experienced less welfare from a given sum of money than the poor did.

In the mid-thirties Lord Lionel Robbins exploded this argument in his *Essay on the Nature and Significance of Economic Science*. He argued that the proposition that one person experiences the same utility as another from a given amount of wealth was scientifically meaningless because it could not be tested.<sup>156</sup> In a widely quoted passage Robbins reasoned

suppose that we differed about the satisfaction derived by A from an income of 1,000 £, and the satisfaction derived by B from an income of twice that magnitude. Asking them would provide no solution. Supposing they differed. A might urge that he had more satisfaction than B at the margin. While B might urge that, on the contrary, he had more satisfaction than A. We do not need to be slavish behaviourists to realise that here is no scientific evidence. There is no means of testing the magnitude of A's satisfaction as compared with B's. If we tested the state of their blood-streams, that would be a test of blood, not satisfaction. Introspection does not enable A to discover what is going on in B's mind, nor B to discover what is going on in A's. There is no way of comparing the satisfactions of different people.<sup>157</sup>

Everything that Robbins said in this passage is very likely true. Nevertheless it did not entirely end the debate over interpersonal utility comparisons. First, Robbins was writing in a positivistic tradition obsessed with limitations on scientific inquiry into state of mind.<sup>158</sup> He simply assumed that the only “scientific” economic inquiry into welfare accordingly concerned

---

<sup>153</sup>ALFRED MARSHALL PRINCIPLES OF ECONOMICS 206 (3d ed. 1895.)

<sup>154</sup>ARTHUR C. PIGOU, THE ECONOMICS OF WELFARE 89 (4<sup>th</sup> ed. 1932).

<sup>155</sup>Jacob Viner, *The Utility Concept in Value Theory and Its Critics* (pts. 1 & 2), 33 J. POL. ECON. 369, 638, 644 (1925). FRANK WILLIAM TAUSSIG, PRINCIPLES OF ECONOMICS 132 (3d ed. 1921) (“[I]nequality of incomes brings a less sum of human happiness than equality of incomes.”); John Bates Clark, *The Ultimate Standard of Value*, 1 YALE REV. 258, 258 (1893) (discussing the social nature of value); Simon N. Patten, *The Scope of Political Economy*, 2 YALE REV. 264, 265 (1894) (similar).

<sup>156</sup>LIONEL ROBBINS, AN ESSAY ON THE NATURE AND SIGNIFICANCE OF ECONOMIC SCIENCE (2d ed. 1935). See HOVENKAMP, OPENING, *supra* note \_\_ at 111-112.

<sup>157</sup>ROBBINS, ESSAY, *supra* note \_\_ at 139-140.

<sup>158</sup>HOVENKAMP OPENING, *supra* note \_\_ at 111-112.

mental state. However, if welfare is estimated objectively, by looking at such external factors as material wealth, health, education, nutrition, shelter, productivity, or numerous other external indicia, then interpersonal comparability is readily possible. As a matter of government policy objective welfare judgments may be superior in any event because they more readily permit generalization about large numbers.<sup>159</sup>

Second, Robbins' analysis overlooked the relationship between wealth of various kinds and human *production* functions. If a transfer from a rich person to a poor person also increases productivity by taking wealth from a person for whom it has little marginal impact and giving it to someone who becomes much more productive, then the transfer may increase welfare by increasing economic growth.

One effect of the welfarism debate was that for some time it created a sharp divide between neoclassical welfare economics and more applied branches of economics, including state economic policy making. The same issue also divided neoclassical welfare economics from most of the other social sciences.<sup>160</sup> In neoclassical economics textbooks, authors talked about wealth transfers as if economics should be indifferent to them. The real concern of economics was said to be allocative efficiency, typically measured by Paretian criteria and the deadweight loss caused by monopoly. By contrast, economic policy during the New Deal and after took the distribution of wealth very seriously and developed important governmental programs designed to give effect to these concerns.<sup>161</sup> Even in public utility law, the goal of universal service for such things as electricity and telephone might be viewed as distributive, at least for those customers who cannot even pay the variable costs of their service.<sup>162</sup>

Today, it seems safe to say, concerns about distribution have dramatically re-entered the picture, although the concerns tend to focus on productivity rather than on Robbins' states of mind.<sup>163</sup>

---

<sup>159</sup> For a more or less contemporary critic on this point, see IAN M.D. LITTLE, A CRITIQUE OF WELFARE ECONOMICS (1950).

<sup>160</sup> See Herbert Hovenkamp, *Knowledge About Welfare: Legal Realism and the Separation of Law and Economics*, 84 MINN. L. REV. 805 (2000).

<sup>161</sup> See HOVENKAMP, OPENING, *supra* note \_\_ at 112-114.

<sup>162</sup> On the development of the concept, see Herbert S. Dordick, *The Origins of Universal Service: History as a Determinant of Telecommunications Policy*, 14 TELECOMMUNICATIONS POL'Y 223 (1990). On telephone service at prices lower than variable costs, see Milton Mueller, *Universal Service in Telephone History: A Reconstruction*, 18 TELECOMM. POL'Y 352 (1993).

<sup>163</sup> E.g., INEQUALITY AND GROWTH: PATTERNS AND POLICY (Kaushik Basu and Joseph E. Stiglitz, eds., 2016); JONATHAN D. OSTRY, ET AL., REDISTRIBUTION, INEQUALITY AND GROWTH (IMF, 2014); Federico Cingano, *Trends in Income Inequality and its Impact on Economic Growth* (OECD working paper, 2014), available at <https://www.oecd-ilibrary.org/docserver/5jxrjncwxv6j-en.pdf?expires=1526298689&id=id&accname=guest&checksum=237F380132F4FA78141554ECACA903DD>; Sutirtha Bagchi & Jan Svejnar, *Does Wealth Inequality Matter for Growth? The Effect of Billionaire Wealth, Income Distribution, and Poverty*, 43 J. COMPARATIVE ECON. 505 (2015)

## Actuarial Conceptions of Value and Risk

The marginalist revolution in economics gave us two fundamental shifts in thinking about economic value. One was the relationship between schedules of demand and the various types of costs that determine output and price. This shift accounts for previously discussed issues such as cost-of-service ratemaking.<sup>164</sup>

The other shift was in forward rather than backward looking conceptions of value. That shift brought much needed realism to the commercial and legal system. For example, the classical notion that the value of a corporation is determined by the amount of historically contributed capital bore no useful relationship to the value of a firm in the eyes of prospective shareholders or creditors.<sup>165</sup> Indeed, the only date on which a firm could meaningfully be said to be worth its paid in value was the day it commenced operations. After that its value could go either up or down, and often in a relatively short period of time. Some firms that started out with little capital became enormous and extravagantly wealthy. Others, which were much better financed to begin with, soon went broke. It all depended on luck, managerial success, the market success of new products or services, and the comparative success or failures of rivals.

All of this is to say that while the marginalist theory of corporate valuation was a significant improvement, reflecting market realities, it enormously complicated decisions about valuation. In classical corporate finance theory a judge could determine whether stock was “watered” – i.e., whether its capital was overstated – by doing a little arithmetic with cash payments and slightly more complex valuation with noncash contributions. But basing the value of a firm on its market prospects was much more difficult, because it included risk and uncertainty in the calculation,<sup>166</sup> as well as a far greater number of variables. The value of a firm became a composite answer to questions about the future value of a firm’s products, its capacity to manufacture, various anticipated headwinds in sources of supply and labor relations, the anticipated health of its rivals, the value and remaining life of intellectual property portfolios, the expected health of the economy overall, and expectations about the national or perhaps world physical and regulatory environment.

In addition, actuarial conceptions of risk, which depend entirely on marginalist tools, facilitated significant changes in the theory of firm risk bearing as well as the insurance industry.<sup>167</sup> The common law itself began to adopt more actuarial, or risk driven, conceptions of

---

<sup>164</sup>See discussion *supra*, text at notes \_\_.

<sup>165</sup>See discussion *supra*, text at notes \_\_.

<sup>166</sup>Under Frank Knight’s terminology a “risk” occurs when the occurrence of a future event is unknown but can be calculated as a probability, such as the future toss of dice. In contrast, “uncertainty” deals with the occurrence of future events whose probability cannot be readily calculated. See FRANK H. KNIGHT, *RISK, UNCERTAINTY, AND PROFIT* (1921); Milton Friedman & L.J. Savage, *The Utility Analysis of Choices Involving Risk*, 56 J. POL. ECON. 279 (1948).

<sup>167</sup>A formative text was ALLAN H. WILLET, *ECONOMIC THEORY OF RISK AND INSURANCE* (1901). The Casualty Actuarial Society was created in 1914 as a professional association of insurance actuaries. Then Professor, later Justice, William O. Douglas became one of the first legal scholars to write about the administration of risk. See William O. Douglas, *Vicarious Liability and Administration of Risk, I & II*, 38 YALE LJ. 584, 720 (1929).

legal duties that governed things such as long-term (relational) contracting and the implementation of negligence and products liability rules in tort.<sup>168</sup>

These developments began early in the history of marginalist thought and have had staggering implications for regulation. The science of risk management has essentially developed into economically sophisticated and technical private and public branches. Private risk management refers mainly to how firms make longer run investment or management decisions in the presence of risk. The public branch is devoted mainly to problems of externalities or spillovers where government intervention is thought necessary.<sup>169</sup> These are most likely to occur when one person's or firm's activities cause harm to someone else.

The balance of this paper focuses on two areas of interest, although there are many, many others.

### *Marginalism and the Common Law from Holmes to Torts*

The first marginalist legal scholar in the United States was Oliver Wendell Holmes, Jr., who devoted much of his academic writing to risk management and the common law legal system. Much has been written about the influence of Darwin on Holmes's thought, but much less about Holmes's marginalism, the other Victorian science that certainly had a more significant impact on his thinking.<sup>170</sup> This was already reflected in Holmes' lectures published as *The Common Law* (1881). In his chapter on tort law he wrote that "... the safest way to secure care is to throw the risk upon the person who decides what precautions shall be taken." Further, the risk must generally be assigned to the person in control of outcomes. He also defended aggressive rules for highly dangerous conduct that "throw the risk upon the party pursuing it." In contract law he argued that contracts should be viewed as "the taking of a risk" and a set of bets about the future. Consequential damages were not appropriate "unless the assumption of that risk is to be taken as having fairly entered into the contract."<sup>171</sup> While Holmes was not an economically technical marginalist, many of these ideas eventually worked their way into the law and economics movement.

Relatedly, Holmes's external standard became a way of conforming individual activity to social norms so as to minimize risk. The legal system cannot assess states of mind; it can only

---

<sup>168</sup>See HOVENKAMP, OPENING, *supra* note \_\_\_\_, Ch. 7.

<sup>169</sup> On private, or entrepreneurial risk management, see JAMES LAM, ENTERPRISE RISK MANAGEMENT: FROM INCENTIVES TO CONTROLS (2d Ed. 2014); ROBERT JARROW, THE ECONOMIC FOUNDATIONS OF RISK MANAGEMENT: THEORY, PRACTICE, AND APPLICATIONS (2016) (mainly financial). On management of public risks, see the now archived reports of the Environmental Protection Agency, Presidential Commission on Risk Assessment and Risk Management, available at <https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=55006>. The Federal Aviation Administration, which is in the Department of Transportation, also maintains a website providing information about its regulatory activities in the area of aviation safety. See [https://www.faa.gov/regulations\\_policies/handbooks\\_manuals/aviation/risk\\_management/](https://www.faa.gov/regulations_policies/handbooks_manuals/aviation/risk_management/)

<sup>170</sup> On Holmes's marginalism, see HOVENKAMP, OPENING at 38-41.

<sup>171</sup>OLIVER WENDELL HOLMES, JR., THE COMMON LAW, 117, 149, 300-305 (1881).

control behavior.<sup>172</sup> For that, Holmes posited the “average” person in “temperament, intellect and education”<sup>173</sup> as the common law norm. Conduct that fell below that norm could rightfully be condemned as negligent.<sup>174</sup>

What Holmes did not fully articulate is that he was turning the “private” common law into a social control device. Holmes worked out the skeleton of a system in which deterrence at the margin became the goal of judge made legal policy, with the “average man” as the determinant of the standard. Further, it was forward looking. Holmes repeatedly emphasized that the operative characteristic of his average person was reasonable foresight. “If the intervening events are of such a kind that no foresight could have been expected to look out for them, the defendant is not to blame for having failed to do so,” Holmes wrote about proximate cause in tort law.<sup>175</sup> Judge Learned Hand’s objective cost-benefit test for negligence augmented this by adding marginalist quantification to foresight.<sup>176</sup> That test – that an action is negligent if the marginal cost of the untaken precaution would have been less than the marginal cost of the accident it would have prevented (probability X magnitude) -- represented the triumph of pure marginalism in tort law.<sup>177</sup> As one pair of authors articulate the test:

The marginal Hand rule states that the injurer is negligent if the marginal cost of his or her precaution is less than the resulting marginal benefit. Thus, the injurer is liable under the Hand rule when further precaution is cost-justified.<sup>178</sup>

Today it has become conventional to regard the *Carroll Towing* rule as requiring a comparison of the marginal cost of precautions against the anticipated cost of an accident. An efficient actor, at the margin, equates the two.<sup>179</sup> That approach has also been substantially

---

<sup>172</sup>*E.g.*, *id.* at 49 (even criminal law is indifferent to states of mind; rather its purpose is “to induce external conformity to rule”); *id.* at 62-63 (“The charge of malice aforethought in an indictment for murder has been shown not to mean a state of the defendant’s mind, as is often thought, except in the sense that he knew circumstances which did in fact make his conduct dangerous.”); *id.* at 136 (“recklessly” does not mean “actual personal indifference to the truth....” Rather, “its means only that the data for the statement were so far insufficient that a prudent man could not have made it without leading to the inference that he was indifferent.... [I]f a man makes his statement on those data, he is liable, whatever was the state of his mind.”).

<sup>173</sup>*Id.* at 63, 108, 303.

<sup>174</sup>*Id.* at 112 (“...when the question of the defendant’s negligence is left to a jury, negligence does not mean the actual state of the defendant’s mind, but a failure to act as a prudent man of average intelligence would have done....”).

<sup>175</sup>*Id.* at 92. See Nicholas St. John Green, *Proximate and Remote Cause*, 4 AM. L. REV. 201 (1870). See HOVENKAMP, OPENING, *supra* note \_\_, at Ch. 7.

<sup>176</sup> *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947). See Chapter 7.

<sup>177</sup>On the commonly accepted view that the relevant costs are marginal, see WILLIAM M. LANDES AND RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF TORT LAW* 87 (1987).

<sup>178</sup>ROBERT COOTER AND THOMAS ULEN, *LAW AND ECONOMICS* 214-215 (2011).

<sup>179</sup>*See, e.g.*, WILLIAM M. LANDES AND RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF TORT LAW* 87 (1987); HOVENKAMP, OPENING, *supra* note \_\_; Allan M. Feldman and Jeonghyun Kim, *The Hand Rule and United States v. Carrol Towing Co. Reconsidered*, 7 Am. L. Econ. Rev. 523 (2005); Robert Cooter, *Unity in Tort, Contract, and Property: the Model of Precaution*, 73 CALIF. L. REV. 1 (1985).



written into the test for dangerous and defective products given in the *Restatement (Third) of Products Liability* that “... a product is defective in design when the foreseeable risks of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design.”<sup>180</sup> This language effectively turns the definition of a dangerous and defective products into an exercise in risk management and cost-benefit analysis. As a common law rule, of course, it does not “regulate” in the command and control sense, but rather imposes liability after the fact. However, federal agencies such as the Consumer Product Safety Commission (CPSC) operate under the same standard for ex ante product safety requirements.<sup>181</sup> The more salient fact is that this rule, much like the previous risk-utility test used in tort cases, incorporates regulatory principles into the common law using marginalist formulations that compare the marginal cost of precautions against marginal benefits – in the words of one important scholar of products liability, “the level at which the marginal cost of the investment equals the marginal cost of product-related accidents thereby avoided.”<sup>182</sup>

### *Innovation Risk and PTAB Patent Review*

Innovation often requires firms to take on a great deal of cost, risk, and uncertainty. A well designed patent system should manage it effectively. The consensus is strong that a patent system is necessary to facilitate innovation,<sup>183</sup> and of course the Constitution expressly authorizes its creation.<sup>184</sup> The appropriate question is not whether to regulate, but rather how to regulate. If the marginal exclusion created by the patent system is too small, some socially valuable inventions will not be created, or perhaps will not be created as early. On the other hand, if the patent system over-excludes, it acts as a clog on the flow of technology. The trick is to find the sweet spot between excessive and inadequate protection.<sup>185</sup>

Considered in isolation, the initial system under which patent applications are evaluated and granted suggests significant deficiencies. First, it is largely *ex parte*, which means that the examiner hears almost exclusively from proponents of the patent.<sup>186</sup> Second, patent examiners

---

<sup>180</sup>RESTATEMENT (THIRD) TORTS: PRODUCT LIABILITY §2(b) (1988). See Aaron D. Twerski and James A. Henderson, Jr., *Manufacturers’ Liability for Defective Product Designs: the Triumph of Risk-Utility*, 74 BROOK. L. REV. 1061 (2009).

<sup>181</sup>See 15 U.S.C. § 2058(f)(2) (2012) (prohibiting the CPSC from promulgating a product safety rule ““unless it has prepared . . . a final regulatory analysis of the rule containing . . . [a] description of the potential benefits and potential costs of the rule....”). See *Zen Magnets, LLC v. Consumer Product Safety Commission*, 841 F.3d 1141 (10<sup>th</sup> Cir. 2016) (disapproving a CPSC rule where cost-benefit analysis was unsupported).

<sup>182</sup>James A. Henderson, Jr., *Product Liability and the Passage of Time: the Imprisonment of Corporate Rationality*, 58 N.Y.U. L. REV. 765, 765 (1983).

<sup>183</sup>Strong, but not unanimous. For one of the exceptions, see MICHELE BOLDRIN & DAVID K. LEVINE, *AGAINST INTELLECTUAL MONOPOLY* (2010) (arguing that patent and copyright be largely abolished).

<sup>184</sup>U.S. Const., Art. I, §8, cl. 8.

<sup>185</sup>See CHRISTINA BOHANNAN AND HERBERT HOVENKAMP, *CREATION WITHOUT RESTRAINT: PROMOTING LIBERTY AND RIVALRY IN INNOVATION*, Chs. 2-3 (2015).

<sup>186</sup>See, e.g., Lee Petherbridge et al., *The Federal Circuit and Inequitable Conduct: An Empirical Assessment*, 84 S. CAL. L. REV. 1293, 1296 (2011) (linking ex parte nature of patent prosecution to need for an inequitable conduct defense).

do not have the time or resources to give each patent application adequate consideration. On average a patent receives eighteen or fewer hours of examination.<sup>187</sup> The average number of claims in a patent hovers at around 16-18,<sup>188</sup> which suggests that examiners spend about an hour per claim evaluating patents, disregarding the time they need to spend on other portions of the patent application. As a result too many invalid patents are granted. When patent validity is litigated in an adversarial setting, where considerably more resources are put into the analysis, as many as 60% or even more of these already issued patents are found completely or partially invalid.<sup>189</sup>

Invalidity rates are only a part of the problem. In addition, a very large number of *legally* valid patents have no economic value. They may be perfectly valid as a matter of patent law, but they are worthless because there is no market for them. The technologies they invent may be too costly in relation to what they provide. There may be better methods for doing the same thing. They may be solutions in search of a problem. In any event, more than 90% of issued patents are never licensed, and of these a high percentage are never commercialized at all.<sup>190</sup> Only 1-3% of issued patents are ever litigated, and the rate has been declining, which is consistent with the proposition that the commercial value of issued patents (aside from validity questions) is declining as well.<sup>191</sup> Further, patent litigation is very expensive, ranging from \$600k where less than \$1m is at risk, to approximately \$5m when the amount at risk exceeds \$25m.<sup>192</sup>

---

<sup>187</sup>See John R. Allison & Mark A. Lemley, *The Growing Complexity of the United States Patent System*, 82 B.U. L. REV. 77, 135 (2002); Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495, 1500 (2001). See also Janet Freilich, *Patent Clutter*, 103 IOWA L. REV. 925 (2018); Roger Allan Ford, *The Patent Spiral*, 164 U. PA., L. REV. 827 (2016); Brenda Sandburg, *Speed Over Substance?*, INTELL. PROP. MAG., Mar. 1999; John R. Thomas, *Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties*, 2001 U. ILL. L. REV. 305, 316-22. And see Michael D. Frakes & Melissa F. Wasserman, *Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Microlevel Application Data*, 99 REV. ECON. & STAT. 550 & app. tbl.A1 (2017).

<sup>188</sup>For recent data, see <https://patentlyo.com/patent/2017/10/standard-patent-size.html>.

<sup>189</sup>See Ronald J. Mann, *A New Look at Patent Quality: Relating Patent Prosecution to Validity*, 9 J. EMPIRICAL LEGAL STUD. 1 (2012) (looking at Federal Circuit decisions, 59.8% found invalid). Other studies largely in accord are John R. Allison, Mark A. Lemley, & David L. Schwartz, *Understanding the Realities of Modern Patent Litigation*, 92 TEX. L. REV. 1769, 1777-1779 (2014). This study updates John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185, 188-93 (1998). See also Paul M. Janicke & Lilan Ren, *Who Wins Patent Infringement Cases*, 34 AIPLA Q.J. 1 (2006) (breaking down data by technology, identify of inventors and accused infringers and law firms). See also Mark A. Lemley, *The Surprising Resilience of the Patent System*, 95 TEX. L. REV. 1 (2016).

<sup>190</sup>See Daniel Fisher, *The Real Patent Crisis is Stifling Innovation*, FORBES (Jun 18, 2014). For a more detailed breakdown of the data, see Ron D. Katznelson, *A Century of Patent Litigation in Perspective* (SSRN, rev. July 31, 2018), available at [https://papers.ssrn.com/sol3/Papers.cfm?abstract\\_id=2503140](https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2503140).

<sup>191</sup>See Katznelson, *id.* And see Hannibal Travis, *Counter-IP Conspiracies: Patent Alienability and the Sherman Antitrust Act*, 71 UNIV. MIAMI L. REV. 758, 762 & . 13 (2017).

<sup>192</sup> See AIPLA Survey of Costs of Patent Litigation and Inter Partes Review (Jan. 30, 2017) (summarized in PatentAttorney.com, available at <https://www.patentattorney.com/aipla-survey-of-costs-of-patent-litigation-and-inter-partes-review/>).

While patent examiners are trained to consider questions of legal validity, they do not generally consider commercial value or success, except in obvious cases of patents that clearly do not work or are useless on their face.<sup>193</sup> One can only imagine, but requiring examiners to test unissued patents for commercial value would increase examination costs enormously, even assuming they were capable of doing it. The best way to test commercial, as opposed to legal, value is through the market. A legally valid patent is commercially valuable if at least one firm wants to license it or produce technology or processes that infringe it.

A rational system for optimizing patent issuance must be responsive to concerns about both invalidity and economic value. Further, it must make these evaluations at reasonable cost. The patent examining process assesses legal validity, but economic value is best assessed by the relevant parties' willingness to put resources into validity determinations. If a patent is economically worthless it is unlikely that people will object to it or dispute its validity, because it does not interfere with their business.

In its 2018 *Oil States* decision the Supreme Court approved a sensible regulatory approach to the system that assesses the legal validity of patents prior to litigation.<sup>194</sup> The challenged administrative process, called inter partes review,<sup>195</sup> permits any person to ask the United States Patent and Trademark Office (USPTO) to reconsider and, if necessary, cancel one or more claims in an issued patent.<sup>196</sup> The procedure is administrative but adversarial. Both the patentee and the challenger have participation rights, including the right to present evidence, make arguments, and have a hearing before the Patent Trial and Appeal Board (PTAB).<sup>197</sup> This makes the process significantly different and more demanding than the type of ex parte examination that goes on during initial patent prosecution. However, it is also much less costly than patent infringement litigation in a federal court. Most particularly, PTAB review is limited to questions of validity under §§102 and 103 of the Patent Act; it does not consider infringement. The PTAB decision is fully reviewable on appeal to the Federal Circuit in which both the patentee and the challenger are entitled to be parties, as well as PTAB itself.<sup>198</sup>

The PTAB proceeding permits the Agency to do a much more thorough and yet relatively low cost review of patent validity.<sup>199</sup> The costs of administrative inter partes review are dramatically lower than the cost of litigation, running around \$275k through a PTAB hearing and \$350k through appeal.<sup>200</sup> One Amicus in *Oil States* estimated the cost differential between

---

<sup>193</sup> For good perspective, see Michael Risch, *Reinventing Usefulness*, 2010 B.Y.U. L. Rev. 1195; Sean B. Seymore, *Making Patents Useful*, 98 MINN. L. REV. 1046 (2014).

<sup>194</sup> *Oil States Energy Svces, LLC v. Greene's Energy Group, LLC*, 138 S.Ct. 1365 (2018).

<sup>195</sup> Created in the Leahy-Smith America Invents Act, 35 U.S.C. §100 et seq. See *id.* at 1371.

<sup>196</sup> 35 U.S.C. §§311-319.

<sup>197</sup> The process is briefly laid out in *Oil States*, 138 S.Ct. at 1371-1372.

<sup>198</sup> See 35 U.S.C. §319; and *Oil States*, 138 S.Ct. at \_\_\_ (describing the appeals procedure).

<sup>199</sup> *Accord* Carl Shapiro, *Patent Reform: Aligning Reward and Contribution*, in 7 INNOVATION POLICY AND THE ECONOMY (Adam B. Jaffe et al. eds., 2007). See also Joseph Farrell and Carl Shapiro, *How Strong are Weak Patents*, 98 AM. ECON. REV. 1347 (2008) (arguing that it would be efficient to create a process that determines patent validity prior to licensing).

<sup>200</sup> *AIPLA Survey*, *supra* note \_\_.

litigation and inter partes review as exceeding ten to one.<sup>201</sup> Importantly, however, the costs are not zero, and as a result someone must have a financial incentive to challenge a patent. Most likely, this is someone producing technology that arguably infringes the patent in question, assuming it is valid. That is to say, patents that are challenged in a PTAB inter partes proceeding very likely do have positive *commercial* value, assuming they are valid, or no one would bother to challenge them.

Some critics of this administrative inter partes review system lament that it undermines a “strong patent system.”<sup>202</sup> That is an odd use of terminology, in which “strong patent system” actually becomes a synonym for protection of weak patents. It is as if a military general who, upon being instructed to raise a strong army, responded by abolishing or downgrading requirements for physical and mental fitness. That might produce a lot of soldiers, but not likely a stronger army. You don’t protect a strong patent system by approving lots of weak patents, but rather by coming up with a rational, cost-effective system for distinguishing those patents that make the requisite contribution to technological progress from those that do not.

The two stage patent evaluation system that inter partes review permits is an efficient application of the economics of information. In George Stigler’s words, “the optimum amount of search will be such that the marginal cost of search equals the expected increase in receipts....”<sup>203</sup> At the initial grant stage relatively few resources are committed to establishing validity – a rational decision given that so many patents are commercially worthless anyway. While the error rate is high, the likelihood that anyone will be harmed is relatively low.<sup>204</sup> Once their commercial value has been declared via a PTAB validity challenge, however, the stakes have gone up and a more costly examination is called for. One additional likely effect is that in the residual patent infringement cases in the district courts a relatively higher percentage of resources will go to questions of infringement, including claim construction, rather than validity. Questions of infringement typically do not second guess the patent issuance process. Rather, they accept the patentee’s claims as given and consider the scope of the accused device or process.

---

<sup>201</sup>Brief for Apple, Inc., as Amicus Curiae in Support of Respondents, *Oil States*, 2017 WL 4946906 (Oct. 30, 2017)

<sup>202</sup> *E.g.*, Brief of University of New Mexico as Amicus Curiae in Support of Petitioner, *Oil States*, 2017 WL 3888211 (U.S.S.Ct. Aug. 31, 2017) (equating elimination of inter partes review with protection of a “strong patent system”); James Carmichael & Brad Close, *Despite Oil States, Inter Partes Review May Still be Held Unconstitutional*, IPWatchdog (April 25, 2018), available at <https://www.ipwatchdog.com/2018/04/25/despite-oil-states-inter-partes-review-may-still-be-held-unconstitutional/id=96406/> (similar) Gene Quinn, *Predicting Oil States in Advance of SCOTUS Oral Arguments*, IPWatchdog (Nov. 12, 2017) (identifying abolition of inter partes review as protecting a “strong patent system”), available at <http://www.ipwatchdog.com/2017/11/12/predicting-oil-states-advance-scotus-oral-arguments/id=89661/> (similar).

<sup>203</sup>George J. Stigler, *The Economics of Information*, 69 J. POL. ECON. 213, 216 (1961).

<sup>204</sup>*Cf.* Lemley, *Rational Ignorance*, *supra* note \_\_.

The patent system, including inter partes review before PTAB, is intended to manage both the very considerable social risk and cost attending innovation as well as its private costs. As with so many marginalist institutions, its job is to thread the needle between a policy that properly incentivizes useful and worthwhile technology while leaving the channels open for other development. An overly aggressive patent system is bad for the dissemination of technological progress just as much as an underdeterrent one.

### **Conclusion**

In our current era in which agency review is under attack it is worth considering exactly what we would have to do to roll the clock back. While this paper does not prove the point, it does suggest that returning to a world in which generalist judges are responsible for devising and executing regulatory policy will require nothing less than reversing the marginalist revolution itself.