THE OLD NATURAL MONOPOLY SOLUTION TO THE NEW BIG TECH PROBLEM

Tyler McNish[*]

ABSTRACT

Is there such a thing as a "good monopoly"? In the 20th century, many people thought so. The concept of natural monopoly posited that certain goods and services were better produced by a regulated monopoly than by a competitive market. Today, however, almost no one uses this idea. We have not identified any new technology as a natural monopoly since the 1930s, and the concept is conspicuously absent from contemporary work on problems of industrial organization, including "Big Tech" market power. This Article reviews 150 years of intellectual history to explain how natural monopoly disappeared from our worldview. I trace the concept's perceived irrelevance to its vestigial neoclassical features. We continue to define natural monopoly as an exception to the general rule of perfect competition, even though we stopped believing in that rule 100 years ago. In this way, we turned natural monopoly into a unicorn—possible to describe in theory but non-existent in the real world. Drawing on the lessons of the imperfect competition revolution and the deregulatory era of the 1970s–1990s, I propose a five-part test for determining whether a product or service should be regulated as a natural monopoly. I then show how this updated theory can help rationalize the respective roles of various proposed Big Tech policy interventions, including antitrust, vertical separations, nondiscrimination rules, interoperability mandates, and shared governance.

* J.D., University of California School of Law, 2010. The author is grateful to Rebecca Talbott for her helpful feedback on this project.

INTRODUCTION

When we say that something is a "natural monopoly," we mean that it ought to be provided by a single enterprise, not by multiple competitors, and (typically) that the enterprise's pricing and conduct ought to be regulated by a specialist public utility commission.[1] Conceived in the late 19th century to explain why the great technological innovations of the era—railroads, telecommunications, and electric power—did not obey the laws of competition as then understood by economic science,[2] the concept of natural monopoly was a fixture of competition policy discourse for most of the 20th century.[3] As late as the deregulatory era of the 1990s, natural monopoly continued to be widely invoked as the criterion for whether or not an industry should remain subject to public utility regulation.[4]

Today, the concept of natural monopoly is conspicuously absent from contemporary debate about the market power of large internet technology enterprises. The dominant line of thinking understands the "Big Tech Problem" as a problem of anticompetitive conduct, to be solved by antitrust, not regulation.[5] The ongoing actions in which the government has challenged various business practices used by Google, Meta, Apple, and Amazon follow this line of thought.[6] So does much of the leading academic work on the subject. The New Brandeisian movement argues that reforms to antitrust doctrine are needed,[7] while antitrust centrists tend to think that proper

---

1. Richard Posner, *Natural Monopoly and Its Regulation*, 21 STAN. L. REV. 548, 548 (1969) [hereinafter *Natural Monopoly and Its Regulation*].

2. Herbert Hovenkamp, *Regulatory Conflict in the Gilded Age: Federalism and the Railroad Problem*, 97 YALE L.J. 1017, 1017–18 (1988) [hereinafter *Regulatory Conflict*].

3. *See. e.g.*, STEPHEN BREYER, REGULATION AND ITS REFORM 158 (1982); ALFRED E. KAHN, THE ECONOMICS OF REGULATION 11–12 (1988) [hereinafter ECONOMICS OF REGULATION]; CHARLES F. PHILLIPS, JR., THE REGULATION OF PUBLIC UTILITIES 45 (1988).

4. *See, e.g.*, Joseph D. Kearney & Thomas W. Merrill, *The Great Transformation of Regulated Industries Law*, 98 COLUM. L. REV. 1323, 1328 (1998); Christopher S. Yoo, *Deregulation vs. Reregulation of Telecommunications*, 36 J. CORP. L. 847, 849 (2011).

5. *See infra* notes 6–8.

6. Cecilia Kang & David McCabe, *After Google Antitrust Ruling, Here's Where the Other Big Tech Cases Stand*, N.Y. TIMES, https://www.nytimes.com/2024/08/05/technology/antitrust-google-amazon-apple-meta.html (Aug. 6. 2024).

7. Lina M. Khan, *Amazon's Antitrust Paradox*, 126 YALE L.J. 710, 716 (2017) [hereinafter *Amazon's Antitrust Paradox*]; Lina M. Khan, *The Separation of Platforms and Commerce,* 119 COLUM. L. REV. 973, 982 (2019) [hereinafter *The Separation of Platforms and Commerce*]; TIM WU, THE CURSE OF BIGNESS: ANTITRUST IN THE NEW GILDED AGE 14–15 (2018) [hereinafter ANTITRUST IN THE NEW GILDED AGE]; Zephyr Teachout, *Antitrust Law, Freedom, and Human Development,* 41 CARDOZO L. REV. 1081, 1096–97 (2019); JOSEPH FISHKIN & WILLIAM E. FORBATH, THE ANTI-OLIGARCHY CONSTITUTION 216–17 (2022).

enforcement of the doctrine we already have is enough,[8] but both groups of thinkers look to competition for solutions, not regulated monopoly.[9] The leading "Big Tech" legislative proposals are also consistent with this approach: for example, the American Innovation and Choice Online Act (AICOA), which was proposed in Congress but not passed in 2022, would have defined new forms of anticompetitive conduct to be enforced ex post facto in the style of antitrust.[10]

With few but important exceptions,[11] even the writers who do contemplate a regulatory solution to the Big Tech Problem—such as interoperability mandates or line of business restrictions—find little use for the concept of natural monopoly. Instead, they use new concepts to define the set of firms that should be specially regulated, like "dominant digital platforms,"[12] "social infrastructure,"[13] "winner-take-all markets,"[14] and "networks, platforms, and utilities."[15] Indeed, even those who want to revitalize the public utility tradition would cleanse it of its traditional association with the natural monopoly concept, reclaiming the Progressive ethos of public utility regulation but rejecting the notion that natural monopoly should be its target.[16]

---

8. Herbert Hovenkamp, *Antitrust and Platform Monopoly*, 130 YALE L.J. 1952, 1972 (2021) [hereinafter *Antitrust and Platform Monopoly*]; Herbert Hovenkamp & Fiona Scott Morton, *Framing the Chicago School of Antitrust*, 168 U. PA. L. REV. 1843, 1878 (2020) [hereinafter *Framing Chicago*]; Daniel A. Crane, *Antitrust's Unconventional Politics*, 104 VA. L. REV. ONLINE 118, 123 (2018) [hereinafter *Antitrust's Unconventional Politics*].

9. *See supra* notes 7–8.

10. American Innovation and Choice Online Act, S. 2992, 117th Cong. (as reported by S. Comm. on the Judiciary, Mar. 2, 2022).

11. In a recent article, Herbert Hovenkamp treats natural monopoly as the guiding principle for Big Tech regulation and articulates a novel test for determining whether natural monopoly is present. *Antitrust and Platform Monopoly*, *supra* note 8, at 1971–72. My goal in this Article is to justify, amplify, and systematize Hovenkamp's insights, explaining why we need a new test for natural monopoly. I also take a different view about the required elements of natural monopoly and draw different conclusions from its application to contemporary problems. Two other recent works also use natural monopoly as a guide to tech regulation, but do not grapple with the limitations of the theory. *See* Tejas Narechania, *Machine Learning as Natural Monopoly*, 107 IOWA L. REV. 1543 (2022) (arguing that machine learning applications may be natural monopolies); Oren Bracha & Frank Pasquale, *Federal Search Commission? Access, Fairness, and Accountability in the Law of Search*, 93 CORNELL L. REV. 1149 (2008).

12. *Separation of Platforms and Commerce*, *supra* note 7, at 982.

13. BRETT FRISCHMANN, INFRASTRUCTURE: THE SOCIAL VALUE OF SHARED RESOURCES 108 (2012).

14. *Antitrust and Platform Monopoly*, *supra* note 8, at 1970 n.67.

15. MORGAN RICKS ET AL., NETWORKS, PLATFORMS, AND UTILITIES LAW AND POLICY 7 (2022).

16. K. Sabeel Rahman, *The New Utilities: Private Power, Social Infrastructure, and the Revival of the Public Utility Concept*, 39 CARDOZO L. REV. 1621, 1638–39, 1680, 1687–88 (2018) [hereinafter *The New Utilities*] (proposing to "excavate" the "ethos" of public utility regulation, not "mechanically copy and reinstate old models of public utility regulation" such as the "tired, old top-down institutional forms we might associate with early twentieth century rate regulation"); *see* K. Sabeel Rahman,

One thing that is surprising about the disappearance of natural monopoly from contemporary discourse is that the concept has never been officially disavowed as a matter of economic theory.[17] It remains routine for legal academics to recite that "[n]atural monopoly exists when the entire demand for a good or service can be satisfied at lowest cost by one firm," albeit only as a tangential reference offered for the purposes of completeness[18] or in support of an argument against regulation,[19] not as part of the main analysis or program of reform.[20] Natural monopoly has become a unicorn: possible to describe in theory but not something we expect to ever see in the wild.

This Article traces the evolution of our concept of natural monopoly across 150 years of intellectual history, with the goal of answering two questions.[21] First, why do we no longer find the concept useful? Second, what should we do about that? Should we abandon natural monopoly in theory, as we have (mostly) done in practice? Or is the idea worth rehabilitating, and perhaps renovating?

On the first question, my thesis is that our concept of natural monopoly is "too neoclassical." By that I mean that we continue to define natural monopoly as an exception to the hypothetical construct of perfect competition, which makes it largely irrelevant to the imperfectly competitive reality in which we now understand ourselves to live. On the second question, my thesis is that natural monopoly, though rarer and more dangerous to regulate than we once may have thought, still describes a real phenomenon, essential to clear thinking about the Big Tech Problem and other contemporary industrial organization problems. If we free the concept of its vestigial neoclassical aspects and update it to work in a world of

---

Infrastructural Regulation and The New Utilities, 35 YALE J. ON REG. 911, 933, 938 (2018) [hereinafter *Infrastructural Regulation*] (the conventional model of regulating natural monopolies should be replaced with a model regulating "[g]oods exhibiting scale, necessity, vulnerability"); William Boyd, *Public Utility and the Low Carbon Future*, 61 UCLA L. REV. 1614, 1708–10 (2014) (arguing for a "revitalized concept of public utility" that "cannot simply adopt the older concept of public utility" but needs "new ideas and conceptual innovations").

17. *See generally* WILLIAM W. SHARKEY, *Natural Monopoly*, *in* 3 THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS 603 (John Eatwell et al. eds. 1987) (recounting the intellectual history of natural monopoly).

18. RICKS ET AL., *supra* note 15, at 9; Erik Hovenkamp, *The Antitrust Duty to Deal in the Age of Big Tech*, 131 YALE L.J. 1483, 1491 n.35 (2022) [hereinafter *The Antitrust Duty to Deal in the Age of Big Tech*].

19. Howard Shelanski, *Information, Innovation, and Competition Policy for the Internet*, 161 U. PA. L. REV. 1663, 1671, 1675–85 (2013).

20. Lina M. Khan, *The New Brandeis Movement: America's Antimonopoly Debate*, 9 J. EUR. COMPETITION. L. & PRAC. 131, 132 (2018) [hereinafter *The New Brandeis Movement*] (disclaiming the view that "big is bad" and noting the historical use of regulation, but focusing on antitrust and the promotion of competition, not natural monopoly regulation).

21. In this Article, I use "our" in the general sense.

imperfect competition, it can help us better determine which features of the Big Tech landscape are adequately handled by antitrust and which demand special regulation.

This Article consists of three Parts. Part I sets out the intellectual history of the natural monopoly concept from its invention in the late 19th century to its fall from prominence in the middle of the 20th century, with an emphasis on certain implications of the imperfect competition revolution that have sometimes been neglected in existing literature. In its earliest incarnation, natural monopoly stood for the neoclassical hope that the classical liberal order might be saved by acknowledging but cabining the threatening implications of industrial scale.[22] The idea was that a few industries might have economies of productive scale sufficient to make them naturally monopolistic, but most were—or should be—perfectly competitive: that is, divided amongst numerous firms competing vigorously on price. The imperfect competition revolution of the 1930s exploded this understanding of competition.[23] We came to understand that an industrialized consumer economy would feature a smaller number of larger firms offering partially differentiated products, and that competition amongst those firms would be imperfect but nevertheless workable. A corollary of this understanding was the belief that even in highly concentrated industries, it was usually better to look to antitrust to enforce what competition could be found than to award a regulated natural monopoly.[24]

Part II explains how we retained a vestigial and mostly unhelpful concept of natural monopoly long after the imperfect competition revolution. In part because of the economics discipline's continued fascination with the concept of perfect competition, we neglected to update the concept of natural monopoly for an imperfectly competitive reality. Consequently, when we drew on the concept during the deregulatory era of the 1970s–1990s to guide decisions about which public utility functions could be made competitive and which could not, the results were not inspiring.[25] Our anachronistic neoclassical understanding of natural monopoly told us little about the key task at hand, namely to distinguish natural monopoly from imperfect competition. Today, natural monopoly remains "on the books" as a theoretical possibility that might occur in extreme scenarios. But we have come to expect that it will be of little help in diagnosing real-world problems of industrial organization.[26]

22. *See infra* Part I.D–E.
23. *See infra* Part I.F.
24. *See infra* Part I.G.
25. *See infra* Part II.B.
26. *See infra* Part II.A.

The dilapidated contemporary state of natural monopoly theory is unfortunate. After all, the deregulatory experience of the 1990s confirmed that some industrial assets or functions *really are* natural monopolies that will resist even our most creative efforts to make them competitive. If natural monopoly is the problem, antitrust alone will not solve it. In the antitrust tradition, monopoly arises from anticompetitive conduct; there is no antitrust liability without such conduct.[27] But if an enterprise controls a natural monopoly, it will not need anticompetitive conduct to dominate its rivals. In that scenario, we will struggle in vain to articulate generally applicable rules of business conduct that correct the market outcomes we are worried about without also netting other conduct that we think is unproblematic. For example, the current wave of Big Tech antitrust litigation challenges Apple's and Google's decisions about which apps to allow on their app stores.[28] Is that conduct more problematic than Walmart's or Costco's curation of their inventories? If so, why? It is difficult to satisfactorily answer this question without resorting to speculation about the inherently monopolistic tendencies of internet platforms[29]—which is the domain of natural monopoly theory, not antitrust.

Part III sets out an updated definition of natural monopoly that frees it of its neoclassical baggage by incorporating the insights of imperfect competition theory and the experience of public utility regulation during the deregulatory era. This requires four main adjustments.[30] First, the traditional technological basis of natural monopoly—economies of scale in production—must be broadened to incorporate some of the characteristics that are often said to make digital industries exceptional, including the centrifugal force of network effects as well as the offsetting centripetal forces of interoperability and multihoming.[31] Second, natural monopoly is only present where product differentiation is substantially absent.[32] Almost all software has economies of scale that would qualify for natural monopoly treatment under the neoclassical understanding of that term. But most software, like many other products and services in an advanced consumer

27. *United States v. Grinnell Corp.*, 384 U.S. 563, 570–71 (1966) (noting that the offense of monopoly under the Sherman Antitrust Act requires "the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident").

28. Kang & McCabe, s*upra* note 6.

29. *See, e.g.*, Fiona Scott Morton et al., *Equitable Interoperability: The "Supertool" of Digital Platform Governance*, 40 YALE J. ON REGUL. 1013, 1016–19 (2023) [hereinafter *Equitable Interoperability*].

30. *See infra* Part III.A.

31. *See infra* Part III.A.

32. *See infra* Part III.A.

economy, is amenable to product differentiation to a degree that allows workable competition. Third, natural monopoly is best judged at the level of an asset or service, not at the level of an entire firm.[33] The right question is not whether Facebook or Google *is* a natural monopoly, but whether they control natural monopoly assets or offer natural monopoly services. Finally, due to the risk that regulation can entrench inefficiency and delay innovation, humility is in order.[34] We should not rush to regulate natural monopolies. And we should act only when we think the risks of regulation are justified by an equally substantial public interest.

The application of this updated natural monopoly theory to contemporary problems of industrial organization can help us better target some of our most promising proposed regulatory interventions—especially the application of interoperability mandates to "Big Tech" enterprises. These mandates have been widely discussed in the academic literature. They have taken on new importance with the recent district court order in the Google Search antitrust case requiring Google to share certain search index data with competitors.[35] If internet platforms can be made to share certain "back-end" utilities—for example, Google's web indexer, or Facebook's database of self-published content—the result may be the kind of competition we most want: multiple "front-ends" built on top of the shared utilities vying to offer us better content moderation systems at lower attentional prices, understood as fewer ads, less behavioral manipulation, less addictive content, or, at the very least, more choices for how to compensate internet gateways for their services. However, the principles that govern the targeting of interoperability remain undertheorized. If Google develops a valuable web indexing platform, and Tesla develops a valuable electric vehicle chassis, why should the law force Google but not Tesla to open its innovation to competitors? For that matter, why Google's web indexer but not its search algorithms?

An updated theory of natural monopoly supplies the missing criterion for intervention: natural monopoly assets, and only those assets, should be subject to interoperability regimes.[36] It also offers something else that is often missing from the existing literature, which is too sanguine about the feasibility of "light touch" rules: a proven framework for making interoperability work. The natural monopoly asset can be spun off and placed

33. *See infra* Part III.A.

34. *See infra* Part III.A.

35. United States v. Google, No. 1:20-cv-03010-APM, slip op. at 8 (D.D.C. Sept. 2, 2025); s*ee* Plaintiffs' Initial Proposed Final Judgment at 18, United States v. Google, No. 1:20-cv-03010-APM (D.D.C. 2024) (explaining the Department of Justice's proposed sharing remedy, which was partially adopted by the District Court).

36. *See infra* Part III.C.

in a public utility, regulated by a specialist commission according to cost-of-service principles developed over a century of regulation.

I do not mean to deny the well-understood imperfections of this approach. We should not imagine that regulation will succeed in setting the prices charged by a social database utility or web indexer utility at the perfectly competitive level—the fool's errand that consumed too much of the efforts of 20th century public utility regulation.[37] Nor can we avoid the risk that regulation might slow dynamic competition, delaying innovations that might bring us better web indexers or content databases. I contend only that these dangers pale in comparison to the danger of entrusting online content moderation and information prioritization to private oligopolies, or worse, empowering government to regulate online speech. The goal is to use regulated monopoly—in the form of back-end utilities—to promote the kind of competition that will make Big Tech content moderation policies no more threatening than newspapers' editorial decisions, and their partnerships no more threatening than those of any other business.

I. THE BIRTH AND (MOSTLY) DEATH OF NATURAL MONOPOLY, 1875–1950

In this Part, I trace the emergence of our ideas about natural monopoly across nearly a century and situate these ideas within broader developments in economic and legal doctrine. To understand the concept's original purpose, it is essential to appreciate that late 19th and early 20th century intellectuals perceived big business as an existential threat to classical liberalism.[38] Natural monopoly stood for the hope that the threatening aspects of industrial scale could be cabined in an exceptional category of regulated public utilities, allowing decentralized competition to continue to govern the mainstream of economic life. When the imperfect competition revolution of the 1930s taught us to instead make peace with big business, the crisis evaporated, and so did natural monopoly's original purpose.

A. Industrial Scale's Threat to the Classical Liberal Order

Our legal and economic doctrine first encountered industrial scale in the form of the "railway problem."[39] The essence of the problem was that "the laws of competition developed in classical economic theory did not work for

37. *See infra* Part III.D.4.

38. *See infra* Part III.A.

39. HERBERT HOVENKAMP, ENTERPRISE AND AMERICAN LAW 141 (1991) [hereinafter ENTERPRISE AND AMERICAN LAW].

the railroads."[40] In the U.S. and Britain, governments often chartered multiple duplicative railroad companies, and encouraged them to compete.[41] However, competition led to overinvestment and price wars.[42] Sometimes the wars ended in bankruptcy of all but one competitor.[43] In other cases, they ended in the consolidation of all competitors into a single system, either by merger or by pooling into cartels.[44] In all of these scenarios, the industry seemed to evolve inevitably towards an equilibrium that left consumers exposed to monopoly pricing.[45] Special government subsidies were sometimes used to preserve the existence of duplicative competitive lines, but those subsidies, no less than monopoly pricing, had a cost to the public.[46]

Contemporary observers were also fascinated by the unprecedented vertical scope of the railroads, which were the first big businesses.[47] Other 19th century transportation systems consisted of public infrastructure and private carriers that operated on that infrastructure. Corporations organized for public benefit often owned highways, waterways, and bridges pursuant to charters that granted monopoly rights and mandated service on specified rates and terms.[48] The diverse individuals, common carriers, and other traffic that operated on this infrastructure were generally competitive and lightly regulated.[49] Early attempts were made to extend this same public-private architecture to railroads by encouraging multiple competitive rail companies to share the same rail lines, but these schemes ran into difficult coordination problems.[50] The complexity of railroad service appeared to demand the vertical integration of the road and the carriers into a monstrous hybrid that extended private control to traditionally public functions.[51]

Around the turn of the 20th century, the railway problem metastasized into the economy-wide "trust problem." Throughout much of the economy, large rationally managed enterprises displaced market transactions among

---

40. *Id.* at 141; *see* GEORGE H. MILLER, RAILROADS AND THE GRANGER LAWS 3 (1971).

41. ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 144; *Regulatory Conflict*, *supra* note 2, at 1031, 1038.

42. ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 144.

43. *Id.* at 148.

44. ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 145.

45. *See generally* CHARLES FRANCIS ADAMS, JR., RAILROADS: THEIR ORIGIN AND THEIR PROBLEMS 81 (N.Y., G.P. Putnam's Sons 1878) (discussing the history of railroad conflict and consolidation); THOMAS K. MCCRAW, PROPHETS OF REGULATION 6, 9–10, 56 (1984).

46. ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 144–47.

47. MILLER, *supra* note 40, at 3.

48. *Id.*

49. *Id.* at 28, 31–32.

50. *Id.* at 27; *Regulatory Conflict*, *supra* note 2, at 1044.

51. *See* MILLER, *supra* note 40, at 24 ("As one contemporary student of the rate problem observed, when a highway is no wider than the wheel of the vehicle which moves upon it, a monopoly of trade for one organization is almost inevitable." (internal quotations omitted)).

individual tradesmen.[52] Increasingly, the "visible hand" of management, not the invisible hand of markets, allocated resources—and productivity appeared to increase in proportion to the change.[53]

In the system of classical economics, this tendency toward monopoly was a puzzle. Classical economics was "antimonopolistic with a vengeance," but its target was *de jure* monopoly, not *de facto* monopoly.[54] For example, classical economists opposed the grant of exclusive trade franchises to joint-stock enterprises like the British East India Company, the award of letters patent to favored manufacturers, and the cartelization of commerce by incorporated guilds.[55] In the classical system, the grant of such monopoly franchises interrupts the functioning of the self-regulating price mechanism, reducing the efficiency with which resources are allocated throughout the economy, and leaving buyers exposed to the monopolist's unchecked self-interest, expressed in the form of high prices. Classical economists assumed that, in the absence of government protection, competition would naturally defeat monopoly and "the weakness of collusion" would frustrate cartels, such that *de facto* monopoly, unlike *de jure* monopoly, ought to be transitory and unthreatening.[56] Why, then, did the new industrial monopolies arise without *de jure* monopoly rights and persist for decades?

With the benefit of hindsight, we can see that the emergence of industrial scale exposed a tension that had been latent in classical economic theory since the beginning: the possibility that the most *allocatively efficient* economic system might not be the most *productively efficient* system. Adam Smith's invisible hand analogy is a theory of allocative efficiency.[57] It posits that competitive market prices adjust to provide a set of incentives and disincentives that guide decentralized human actors to efficiently allocate resources.[58] Smith also had a theory of productive efficiency, which he described with an analogy to a pin factory.[59] Smith observed that an

52. The business trust was a Gilded Age innovation used to roll up multiple corporations under common management. By the 1890s, the trust fell out of favor as a legal instrument and was replaced by other forms of organization. ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 244–45.

53. *See* ALFRED D. CHANDLER, JR. THE VISIBLE HAND: THE MANAGERIAL REVOLUTION IN AMERICAN BUSINESS 1, 6 (1977); RICHARD HOFSTADTER, THE AGE OF REFORM 213–27 (1955).

54. *Regulatory Conflict*, *supra* note 2, at 1030.

55. ADAM SMITH, THE WEALTH OF NATIONS 69, 142–43, 485, 693, 814–15 (Edwin Cannan ed., Modern Library 2000) (1776).

56. Herbert Hovenkamp, *The Antitrust Movement and the Rise of Industrial Organization*, 68 TEX. L. REV. 105, 144 (1989) [hereinafter *Antitrust Movement*]; George J. Stigler, *The Economists and the Problem of Monopoly* 5–6 (U. Chi. Law, Occasional Paper No. 19, 1983).

57. MARK BLAUG, ECONOMIC THEORY IN RETROSPECT 57 (5th ed. 1997) [hereinafter ECONOMIC THEORY].

58. *Id.*

59. SMITH, *supra* note 55, at 4.

individual pin maker could make at most only a few pins per day.[60] In a pin factory, by contrast, the "business of making a pin . . . [is] divided into about eighteen distinct operations . . . all performed by distinct hands," and this division of labor results in productive efficiencies that allow each worker to make "forty-eight thousand pins in a day."[61]

Smith thought the invisible hand and pin factory worked in concert, and in his era, they mostly did.[62] In 1776, actual factories with in-firm economies of scale were still a relatively unimportant mode of economic organization.[63] Smith used the division of labor *within* the pin factory merely as a convenient analogy for the division of labor *among* firms throughout the economy.[64]— The latter division of labor was what Smith cared most about, for it drove his theory of economic growth: an expanding market facilitates a finer division of labor amongst individual tradespeople,[65] which increases overall production and prosperity, which further expands the market, and so on, in a positive feedback loop of increasing returns all guided by an invisible hand of price signals.[66] This was the gist of Smith's famous description of an English laborer's woolen coat as "the produce of the joint labour of . . . [t]he shepherd, the sorter of the wool, the wool-comber or carder, the dyer, the scribbler, the spinner, the weaver, the fuller, the dresser," which in turn depend on "[t]he miner, the builder of the furnace for smelting the ore, the feller of the timber," and so on.[67] Smith thought this division of labor explained why the humble English laborer was better clothed than "many an African king."[68]

A century later, when the division of labor *inside* firms became impossible to ignore, the logic of Smith's pin factory came to stand in opposition to the logic of his invisible hand. For if the division of labor increases with the extent of the market, and labor can be more efficiently divided within firms than between them, why not concede the whole market

60. *Id.*

61. *Id.* at 5.

62. *Id.* at 14–15 (explaining that the "The Principle Which Gives Occasion to the Division of Labor" is self-interested competition).

63. ROBERT L. HEILBRONER, THE WORLDLY PHILOSOPHERS: THE LIVES, TIMES AND IDEAS OF THE GREAT ECONOMIC THINKERS 58, 72 (7th ed. 1999).

64. *See* DAVID WARSH, KNOWLEDGE AND THE WEALTH OF NATIONS: A STORY OF ECONOMIC DISCOVERY 38–40 (2006) (explaining the relationship between the pin factory example and Smith's theory of growth).

65. As Smith put it "*The Division of Labor is Limited by the Extent of the Market*." SMITH, *supra* note 55, at 19.

66. *See* WARSH, *supra* note 64, at 38–43 (explaining the relationship between the division of labor and the invisible hand).

67. SMITH, *supra* note 55, at 12.

68. SMITH, *supra* note 55, at 12–13; ECONOMIC THEORY, *supra* note 57, at 35.

to a single firm so that it could maximize the division of labor? Why have Coca-Cola *and* Pepsi-Cola, with duplicative management hierarchies, duplicative marketing budgets, and duplicative bottling plants? Yet if we abandon ourselves to monopoly, we undermine the invisible hand's ability to channel individual decisions under conditions of freedom towards socially productive goals.[69] In this way, the emergence of industrial scale seemed to suggest—contrary to the principles of classical political economy—that we might have to choose between the goals of productive efficiency and the decentralized allocation of resources.[70]

This was not an arcane problem of economic science. It was a threat to the whole 19th century edifice of classical liberalism. American democracy was thought to rest on a political economy of yeoman farmers and small businessmen interacting under conditions of rough equality, but this was threatened by the concentration of economic and political power in large corporations.[71] American morality was thought to require individuals to run the "race of life" according to their own ideas and conscience, but that lifestyle would be unavailable to employees directed by bosses in degradingly hierarchical organizations.[72] American law structured this liberal moral-economic-political order with property rights, freedom of contract, and hostility to "class legislation"[73]—but these doctrines bent sinister when applied to the benefit of corporations instead of individual tradespeople.[74]

Many of the era's intellectuals feared that industrial consolidation led inexorably to socialism. For if the new concentrations of private power could not be checked by the market, they would need to be subjected to the will of the public expressed through a powerful state. Economist John Bates Clark wrote in the Atlantic Monthly in 1900 that the question of whether monopoly

---

69. As George Stigler later put it: "[H]ere was the dilemma: either the division of labor is limited by the extent of the market, and, characteristically, industries are monopolized; or industries are characteristically competitive, and the [invisible hand] theorem is false or of little significance." George J. Stigler, *The Division of Labor is Limited by the Extent of the Market*, 59 J. POL. ECON. 185, 185 (1951). A similar question led Coase to his theory of the firm: if the invisible hand works so well, why do we have firms at all? Ronald H. Coase, *The Nature of the Firm (1937)*, *in* THE NATURE OF THE FIRM: ORIGINS, EVOLUTION, AND DEVELOPMENT 18, 19 (Oliver E. Williamson & Sidney G. Winter eds., 1993).

70. HEILBRONER, *supra* note 63, at 160–61; JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM, AND DEMOCRACY 106 (3d ed. 2010).

71. HOFSTADTER, *supra* note 53, at 23.

72. *Id.* at 223.

73. WILLIAM M. WIECEK, THE LOST WORLD OF CLASSICAL LEGAL THOUGHT: LAW AND IDEOLOGY IN AMERICA, 1886–1937 107, 135 (1998).

74. *See, e.g.*, Lochner v. New York, 198 U.S. 45 (1905).

was essential to modern prosperity was "momentous beyond the power of language to measure"[75]:

> Answer this question in one way, and you will probably be a socialist; and of course you ought to be one. Answer it in another way, and . . . [y]ou will believe in freedom of individual action, in competition, in the right of contract; in short, in the things that have made our civilization what it is.[76]

Similarly, conservative Seventh Circuit Judge Peter Grosscup wrote in 1905, "what shall it profit our country if it gain the world, and lose its soul?"[77] Grosscup observed that:

> The transformation of the ownership of a country's industrial property, from its people generally, to a few of its people only, reaches the bed-rock of social and moral forces on which, alone, the whole structure of republican institutions rests . . . [I]nstead of depending, each on himself and his own intelligence chiefly for success, the great bulk of our people, increasingly, will become dependents upon others.[78]

This meant "social and, eventually, political revolution."[79]

In 1914, economists John Bates Clark and John Maurice Clark put it even more plainly: if the trusts have "come to stay," then society faces "a choice between the devil of private monopoly and the deep sea of state socialism."[80]

B. The Antitrust Tradition vs. the Public Utility Tradition

The threat of big business inspired two divergent intellectual responses.[81] One response was the antitrust tradition, which coalesced in the Sherman Antitrust Act of 1890. In this tradition, big firms are not typically more productive than small ones. Instead, big firms prevail over smaller

---

75. John Bates Clark, *Disarming the Trusts,* ATLANTIC MONTHLY, Jan. 1900, at 49.
76. *Id.*
77. Peter S. Grosscup, *How to Save the Corporation*, MCCLURE'S MAG., Feb. 1905, at 443, 443.
78. *Id.* at 447.
79. *Id.* at 444.
80. JOHN BATES CLARK & JOHN MAURICE CLARK, THE CONTROL OF TRUSTS 1–2 (1914) [hereinafter CONTROL OF TRUSTS 1914].
81. *Id.*at 141.

rivals by using their economic and political power to forestall competition and dominate markets. Law should forbid such anticompetitive conduct in order to revitalize competition and renew the economic foundations of American democracy and moral society.

The second response was the public utility tradition, which emerged out of the railroad regulatory commissions first set up by state legislatures in the 1870s.[82] In this tradition, big firms typically *are* more productively efficient than small firms, because they benefit from economies of scale. The yearning to reinstate atomistic competition is therefore atavistic and quixotic. Law should embrace but regulate monopoly, so that society may benefit from productive efficiencies without being exposed to the high prices, discrimination, and other depredations it might otherwise impose.

For half a century between 1890 and 1940, these two fundamental visions of what to do about the problem of economic scale clashed repeatedly across diverse intellectual battlefields.[83] We now tend to understand antitrust and public utility regulation as two different tools for fine-tuning economic performance, appropriate to different circumstances. In an earlier era, however, each of the two traditions were seen as an economy-wide solution to the fundamental problem of industrial scale, meaning that they pointed in starkly different directions. Writing in the first years of the 20th century, John Bates Clark and John Maurice Clark summarized the scene as follows:

> Among those who approach the question fairly and intelligently, there are two kinds of plans proposed, springing from two views of the fundamental nature of the ills that now beset us.
>
> The first, and perhaps most widely held among business men, is that in large-scale business competition has failed completely and monopoly has come to stay. The large plant is more efficient than the small one, the combination is more efficient than the independent, competition is wasteful and unnatural and monopoly the inevitable outcome. . . . If they have their way they will legalize monopoly, and in place of free competition as the regulator of prices, they will place the decrees of a public commission.

82. *Id.*; ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 127–130 (the discipline of public utilities law evolved from the Supreme Court's decision in *Munn* to affirm the constitutionality of price regulation by state commission only as applied to "quasi-public" corporations).

83. *See* HOFSTADTER, *supra* note 53 (providing a book-length evocation of this clash and its many ripple effects).

> The other way of attacking the problem starts from a widely different diagnosis. It rests on the belief, deep rooted in the minds off the masses of our people, that competition is not yet dead, that the monopolistic power of the trusts are accidental and not inevitable, that they are built upon privileges that can be removed, powers that can be withdrawn, and predatory acts that can be forbidden. Those who hold such a view naturally wish first to forbid every form of unfair advantage which one competitor may take over his rivals, and further to forbid combination, in whatever guise, when it goes beyond the point at which effective competition can survive.[84]

The 1912 presidential election put the choice between these divergent approaches to the electorate.[85] Reversing his previous reputation as a trustbuster, Teddy Roosevelt ran on a platform of monopolizing industry but subjecting it to political control—a public utility-ization of the whole economy.[86] In his view, "[c]ombinations in industry are the result of an imperative economic law which cannot be repealed by political legislation. . . . The way out lies, not in attempting to prevent such combinations, but in completely controlling them in the interest of the public welfare."[87]

Woodrow Wilson agreed with Roosevelt about the problem, observing that the "one great basic fact which underlies all the questions that are discussed on the political platform at the present moment" was that "the individual has been submerged," such that "men [now] work, not for themselves, not as partners in the old way in which they used to work, but generally as employees . . . of great corporations."[88] In contrast to Roosevelt, however, Wilson ran on a platform of antitrust, proposing to "regulate competition," not "regulate monopoly."[89] Wilson's goal was the restoration of an individualistic economy and society in which "eager men were everywhere captains of industry, not employees; not looking to a distant city to find out what they might do, but looking about among their neighbors,

84. CONTROL OF TRUSTS 1914, *supra* note 80, at 141–43.

85. Daniel A. Crane, *All I Really Need to Know About Antitrust I Learned in 1912*, 100 IOWA L. REV. 2025, 2027 (2015) [hereinafter *All I Really Need to Know*].

86. *Id.*

87. Theodore Roosevelt, The New Nationalism (August 31, 1910).

88. Woodrow Wilson, The New Freedom 3, 5 (1913).

89. *All I Really Need to Know*, *supra* note 85, at 2031. Wilson was advised by Louis Brandeis, then a leading figure in the antitrust movement and a critic of what he called "the curse of bigness." *Id.* at 2028.

finding credit according to their character, not according to their connections."[90]

Wilson won the election, but the debate over what to do about industrial scale was not so finally resolved.[91] During the Great Depression, it sprang again to the foreground. We now understand Depression-era overproduction as a temporary and preventable macroeconomic virus, but at the time it was widely seen as a symptom of the incompatibility of free market capitalism with economies of scale.[92] New Deal economists and reformers were thus absorbed by "the old, scholastic conundrum of 1912," tending to think "in terms of two general solutions, one involving industrial atomization to restore a self-adjusting economy, the other involving centralized planning and detailed regulation."[93] In 1938, a contemporary observer summarized the ambivalence as follows:

> Two souls dwell in the bosom of this Administration, as, indeed, they do in the bosom of the American people. The one loves the Abundant Life, as expressed in the cheap and plentiful products of large-scale mass production and distribution. . . . The other soul yearns for former simplicities, for decentralization, for the interests of the 'little man,' . . . denounces 'monopoly' and 'economic empires,' and seeks means of breaking them up. Our administration manages a remarkable . . . stunt of being . . . in favor of organizing and regulating the Economic Empires to greater and greater efficiency, and of breaking them up as a tribute to perennial American populist feeling.[94]

The seesaw between the First and Second New Deals expressed this ambivalence. The First New Deal followed the public utilities line of thinking: the National Industrial Recovery Act of 1933 (NIRA) attempted the government-organized cartelization of much of the economy, replacing industrial competition with cooperation and scale.[95] When NIRA was invalidated by the Supreme Court, the Second New Deal pivoted to the reinvigoration of antitrust, and a rhetoric focused on reducing scale and

90. HOFSTADTER, *supra* note 53, at 224.

91. *Id.* at 245–56.

92. *See id.* at 302–03; ELLIS W. HAWLEY, THE NEW DEAL AND THE PROBLEM OF MONOPOLY: A STUDY IN ECONOMIC AMBIVALENCE 12–13 (1966).

93. HAWLEY, *supra* note 92, at 187, 420.

94. Dorothy Thompson, *On the Record*, N.Y. HERALD TRIBUNE, Jan. 24, 1938, at 17, 17.

95. HAWLEY, *supra* note 92, at 15.

restoring competition.[96] Even this move, however, was shot through with ambivalence: Thurman Arnold, who Roosevelt appointed to lead the effort, was an antitrust skeptic who had previously argued that "[t]he actual result of the antitrust laws was to promote the growth of great industrial organizations by deflecting the attack on them into purely moral and ceremonial channels."[97] Arnold thought the popular sentiment in favor of antitrust revealed "a society which unconsciously felt the need of great organizations, and at the same time had to deny them a place in the moral and logical ideology of the social structure."[98]

World War II gave this battle of ideas military form. European fascism had taken the path that the U.S. rejected in 1912 and again in the early 1930s, embracing the rationalization of the economy by monopolies and cartels subordinated to the will of the state.[99] The economic results were striking: Hitler's Germany nearly doubled the size of its economy between 1932 and 1939.[100] Similarly, the Soviet Union's rapid industrial development under the Five-Year Plans of the 1920s and 1930s was an economic achievement without precedent in the capitalist economies.[101] Britain and the U.S. carried the flag of free enterprise into World War II, but their victory was procured in large part through wartime industrial planning measures that broke with competitive traditions.[102] As late as 1944, Friedrich Hayek framed his advocacy for free market organization as a voice in the wilderness: he found the English-speaking intelligentsia of his time convinced of the superiority of centralized rationalization of the economy.[103]

---

96. *Id.* at 15, 137–38.

97. THURMAN W. ARNOLD, THE FOLKLORE OF CAPITALISM 212 (1937).

98. *Id.* at 211.

99. Daniel A. Crane, *Fascism and Monopoly*, 118 MICH. L. REV. 1315 (2020) (stating that turn-of-century German law encouraged the rationalization of economic activity by cartels, which banks reorganized as monopolies during the economic crisis of the 1920s, and which then came to form the backbone of the Nazi project of national rearmament in the 1930s).

100. Yuri Kofner, *150 Years of German Monetary History*, INST. FOR MKT. INTEGRATION & ECON. POL'Y (Jan. 4, 2023), https://miwi-institut.de/archives/2626 ("[T]he first half of the Nazi regime, i.e., from 1933 to 1939, was characterized by a strong economic upswing (annual GDP growth rates of over 8 percent), a reduction in unemployment and low consumer price inflation of around 0.7 percent annually.").

101. PETER KENEZ, A HISTORY OF THE SOVIET UNION FROM THE BEGINNING TO THE END 91–92 (1999).

102. *See* ZACHARY D. CARTER, THE PRICE OF PEACE: MONEY, DEMOCRACY, AND THE LIFE OF JOHN MAYNARD KEYNES 353, 410 (2020).

103. FRIEDRICH HAYEK, THE ROAD TO SERFDOM 25 (1944). In fact, these fears persisted in some form well into the 1950s and 1960s. Paul Krugman pointed out that articles in publications such as Foreign Affairs during that era routinely concluded that collectivist, authoritarian states were "inherently better at achieving economic growth than free market democracies." Paul Krugman, *The Myth of Asia's Miracle*, 73 FOREIGN AFFAIRS, Nov.–Dec. 1994, at 62, 65.

C. The Natural Monopoly Synthesis

As originally conceived, the concept of natural monopoly was a kind of synthesis of the antitrust tradition and the public utilities tradition. It stood for the hope that industrial scale could be cabined into a set of exceptional industries—the natural monopolies—such that competition could continue to govern the mainstream of economic life.

The attraction of such an outcome to those who valued the status quo was obvious. Indeed, in *Munn v. Illinois* (1876), the Supreme Court had fashioned a legal doctrine with a similar purpose, a decade or two before the concept of natural monopoly was widely known.[104] The Court held that invasive state economic regulation could be constitutionally applied to an exceptional category of businesses (including railroads) that were "affected by a public interest," even though such regulation would be unconstitutional under the substantive due process principles of *Lochner*[105] if applied to an ordinary business.[106] For the next 60 years, the Court was called on to decide which industries fell within this special category; the ones that did became the public utilities.[107] This *Munn* line of decisions, however, was famously unpersuasive.[108] The Court never satisfactorily reconciled the idea of "affected with a public interest" with the prevailing principles of classical political economy.[109]

Charles Francis Adams, Jr. was one of the first to articulate a more persuasive economic justification for why some industries might require such special regulation.[110] Adams was a scion of the Adams political dynasty,

---

104. Munn v. Illinois, 94 U.S. 113 (1876); For useful commentary on *Munn*, see Harry N. Scheiber, *The Road to Munn: Eminent Domain and the Concept of Public Purpose in the State Courts*, 5 PERSPS. AM. HIST. 329 (1971); William J. Novak, *Law and the Social Control of Capitalism*, 60 EMORY L.J. 376, 402 (2010).

105. Lochner v. New York, 198 U.S. 45 (1905).

106. *See Munn*, 94 U.S. at 126 (quoting Matthew Hale, *De Portibus Maris*, *in* A COLLECTION OF TRACTS RELATIVE TO THE LAW OF ENGLAND 45, 78 (Francis Hargrave ed., 1787).

107. RICHARD F. HIRSH, POWER LOSS: THE ORIGINS OF DEREGULATION AND RESTRUCTURING IN THE AMERICAN ELECTRIC UTILITY SYSTEM 12–22 (1999); BARBARA H. FRIED, THE PROGRESSIVE ASSAULT ON LAISSEZ-FAIRE: ROBERT HALE AND THE FIRST LAW AND ECONOMICS MOVEMENT 11, 96 (1998).

108. Herbert Hovenkamp, *The Political Economy of Substantive Due Process*, 40 STAN. L. REV. 379, 440 (1988); FELIX FRANKFURTER, THE COMMERCE CLAUSE UNDER MARSHALL, TANEY, AND WAITE 87 (1937).

109. FRANKFURTER, *supra* note 108, at 85.

110. MCCRAW, *supra* note 45, at 9–11; Herbert J. Hovenkamp, *The First Great Law & Economics Movement*, 42 STAN. L. REV. 993, 994–97 (1990) [hereinafter *First Great Law & Economics Movement*]. The first glimmerings of the idea appeared in John Stuart Mill's 1848 treatise, but the point was underdeveloped, and it is not clear whether Mill was talking about the same thing as later users of the term "natural monopoly." JOHN STUART MILL, PRINCIPLES OF POLITICAL ECONOMY 391 (London, Parker & Co. 1848).

Civil War hero, halfhearted lawyer, muckraking literary journalist, pioneering public utility regulator, and railroad executive.[111] By the late 1860s, Adams had worked out that railroads had unusually high economies of scale: "[i]t is an undisputed law of railway economics that the cost of the movement is in direct inverse ratio to the amount moved." This implied "a conclusion which is at the basis of the whole transportation problem: competition and the cheapest possible transportation are wholly incompatible."[112] By the 1870s, Adams had worked these ideas up into a fairly complete theory of natural monopoly:

> The traditions of political economy to the contrary notwithstanding, there are functions of modern life, the number of which is also continually increasing, which necessarily partake in their essence of the character of monopolies. This they do and always must do as the fundamental condition of their development. Now it is found that, wherever this characteristic exists, the effect of competition is not to regulate cost or equalize production, but under a greater or less degree of friction to bring about combination and a closer monopoly.[113]

In his 1887 essay *The Relation of the State to Industrial Action*, economist Henry C. Adams (no relation) expanded on these ideas and connected them to the doctrine of classical political economy. Henry Adams identified a "class of industries" (including railroads) which "conform[ed] to the law of increasing, rather than to the law of constant or decreasing returns."[114] He observed that John Stuart Mill—who wrote the century's leading treatise on classical political economy—was aware of the existence of such increasing returns to scale but failed to appreciate the implications of the phenomenon: namely, that "where the law of increasing returns works with any degree of intensity, the principle of free competition is powerless to exercise a healthy regulating influence."[115] In this scenario, Henry Adams thought that "there can be no question as to the line which marks the duties

111. MCCRAW, *supra* note 45, at 2–6.

112. *Id.* at 9 (quoting Charles Francis Adams, Jr., *Railway Commissions*, 2 J. SOC. SCI. 233, 234 (1870) (emphasis omitted)).

113. Charles Francis Adams, Jr., *The State and the Railroads*, ATLANTIC MONTHLY, June 1876, at 691, 692.

114. HENRY C. ADAMS, THE RELATION OF THE STATE TO INDUSTRIAL ACTION 57, 59–61 (Balt., Am. Econ. Ass'n 1887).

115. *Id.* at 60.

of the state. . . . The control of the state over industries should be co-extensive with the application of the law of increasing returns."[116]

Two years later, economist and Progressive reformer Richard T. Ely credited Henry Adams with demonstrating "the impossibility of competition in a business like the telegraph service," and described such enterprises as "natural monopolies."[117] Ely argued that "there is a certain class of pursuits for which there is no escape from monopoly. . . . We consequently see that we have a choice between two alternatives, and there is no middle ground between them. These are (a) private monopoly; and (b) public monopoly . . . ."[118] Ely advocated public monopoly as "the lesser of the two evils."[119]

By the 1890s and early 1900s, the idea of technologically determined natural monopolies that should be subjected to public control was widespread in Progressive literature, with Ely,[120] John Commons,[121] and John Bates Clark[122] all prominently exploring its implications. In an 1894 article, Ely succinctly described what the "natural monopoly" term had come to mean:

> There are various undertakings . . . virtually all of them comparatively new . . . which are monopolies by virtue of their own inherent properties. Recent discussions have made these businesses well known. They are railways, telegraphs, telephones, canals, irrigation works, harbors, gasworks, street-car lines, and the like. Experience and deductive argument alike show that in businesses of this kind there can be no competition, and that all appearances which resemble competition are simply temporary and illusory.[123]

The two Adamses, Ely, and Commons conceived natural monopoly as an apostasy, part of their assault on the *laissez-faire* cathedral of classical economics.[124] Alfred Marshall, however, soon found it convenient to incorporate a similar concept into the mainstream of British political

116. *Id.*
117. Richard T. Ely, *Telegraph Monopoly*, 149 N. AM. REV. 44, 44–45 (1889).
118. *Id.* at 45–46.
119. *Id.* at 47.
120. Richard T. Ely, *Municipal Ownership of Natural Monopolies*, 172 N. AM. REV. 445 (1901).
121. John R. Commons, *Protection and Natural Monopolies*, 6 Q.J. ECON. 479 (1892).
122. JOHN BATES CLARK, THE CONTROL OF TRUSTS 1 (1901).
123. Richard T. Ely, *Natural Monopolies and the Workingman. A Programme of Social Reform* 158 N. AM. REV. 294, 294 (1894).
124. FRIED, *supra* note 107, at 11.

economy.[125] Marshall accomplished this using the tools of the marginal revolution, which marks the break between classical and neoclassical economics. Neoclassical thinkers rebuilt their discipline on the cornerstone of an assumption Marshall called the "law of diminishing return."[126] On the demand side, this law reflects the fact that consumers tend to value their sixth apple less than their first.[127] On the supply side, it reflects the fact that it costs orchard owners less to pick their first ton of apples than to pick their 6,000th ton, which will be eked from less suitable land, less productive trees, or higher branches.[128] Market-wide, the diminishing returns experienced by consumers and producers add up to the downward-sloping demand curve and upward-sloping supply curve of the famous *Marshallian Cross* diagram, depicted in the left panel of Figure 1.

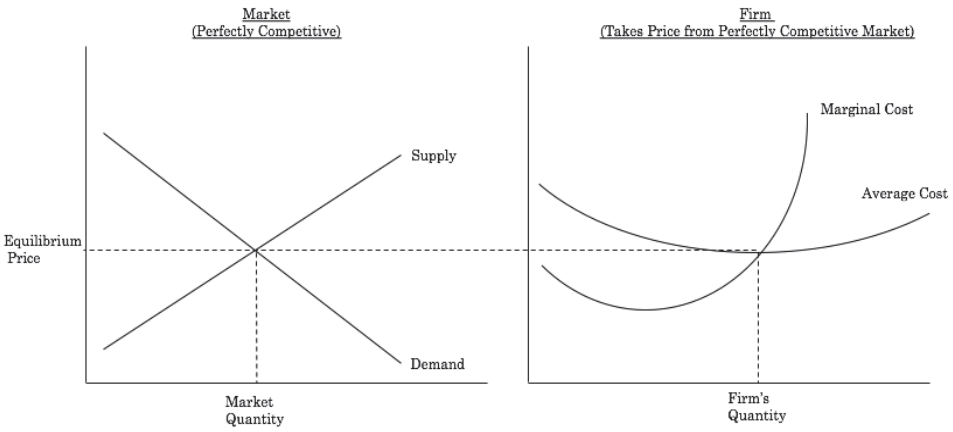

Figure 1: Equilibrium Under Perfect Competition[129]

125. *Infra* notes 122–24 and accompanying text.

126. ALFRED MARSHALL, PRINCIPLES OF ECONOMICS 288 n.1 (8th ed. 1920); ECONOMIC THEORY, *supra* note 57, at 277–78; *see* BEN B. SELIGMAN, MAIN CURRENTS IN MODERN ECONOMICS 461–65 (2d ed. 1990); WARSH, *supra* note 64, at 76–78. Stanley Jevons, Leon Walras, and Carl Menger are generally credited with the key insights that sparked the marginal revolution. Milton Friedman, *Leon Walras and His Economic System*, 45 AM. ECON. REV. 900, 901 (1955); Herbert Hovenkamp, *The Marginalist Revolution in Legal Thought*, 46 VAND. L. REV. 305, 308–13 (1993). Alfred Marshall compiled their insights into his stylish, comprehensive treatise, which I rely on here as the best summary of the neoclassical perspective. MARSHALL, *supra* note 126.

127. MARSHALL, *supra* note 126, at 78–79.

128. *Id.* at 125.

129. ROBERT H. FRANK, MICROECONOMICS AND BEHAVIOR 348 (3d ed. 1997); MARSHALL, *supra* note 126, at 318 n.1.

This elegant model, still the foundation for "almost everything we know about the behavior of the economic system,"[130] reconciles the above-described tension between the pin factory and the invisible hand. In fact, extended to a general equilibrium context—in other words, to an economy containing multiple linked markets, each with its own supply and demand schedule—the neoclassical model of competition can be shown to result not only in an equilibrium, but in a social optimum.[131] In this happy world of perfect competition, the market-clearing price equals the industry's marginal cost of production, which equals the marginal consumer's marginal rate of substitution between the good and all other goods—and, best of all, we can prove this using mathematics.[132]

The result only holds, however, if there is no incompatibility between the optimally productive firm size and the number of firms needed for competition. The neoclassical model denies any such incompatibility by assuming that within a typical firm, increasing returns will be present only up to a certain quantity of production, beyond which they will be dominated by decreasing returns to scale.[133] Thus, a firm's marginal-cost curve may slope down for a certain range (as it did in Smith's pin factory), but must ultimately turn upward, as shown in the right panel of Figure 1.[134] As long as these cost curves start sloping upward at a quantity sufficiently below the total market demand, the necessary output will be produced more efficiently by multiple firms than by one—and these firms will compete.[135] Competition is crucial to the neoclassical model, for it is the mechanism that equilibrates price to the marginal firm's marginal cost of production. A firm that rebels against this market price will merely inspire other firms to ramp up their production, and consumers will purchase these firms' goods, not the rebel firm's overpriced offerings.[136]

---

130. ECONOMIC THEORY, *supra* note 57, at 337.

131. Mark Blaug, *The Fundamental Theorems of Modern Welfare Economics, Historically Contemplated*, 39 HIST. POL. ECON. 185, 185–86 (2007) [hereinafter *Fundamental Theorems*]; ECONOMIC THEORY, *supra* note 57, at 579.

132. John Geanakoplos, *Arrow-Debreu Model of General Equilibrium*, *in* THE NEW PALGRAVE DICTIONARY OF ECONOMICS 116, 118–19 (John Eatwell et al. eds., 1987); ECONOMIC THEORY, *supra* note 57, at 552–53; Kenneth Arrow, *General Economic Equilibrium: Purpose, Analytic Techniques, Collective Choice*, *in* NOBEL LECTURES IN ECONOMIC SCIENCES 1969–1980, at 109, 113–15 (Assar Lindbeck ed., 1992); ECONOMICS OF REGULATION, *supra* note 3, at 65 ("The central policy prescription of microeconomics is the equation of price and marginal cost.").

133. ECONOMIC THEORY, *supra* note 57, at 355–63.

134. *Id.* at 356–57.

135. *Id.* at 362–69.

136. Thus, competitive firms are said to be "price takers": they are forced to accept the market price for their goods. FRANK, *supra* note 129, at 377.

From the beginning, however, neoclassical economists realized that this elegant reconciliation of competition with efficient firm scale did not describe *every* part of the modern economy. Marshall articulated not only a "law of diminishing returns" but also a "law of increasing returns."[137] This second law recognized that "in those industries which are not engaged in raising raw produce an increase of labour and capital generally gives a return increased more than in proportion."[138] In other words, firms in these industries don't have U-shaped marginal cost curves like that shown in Figure 1. Instead, their marginal cost curves continue to slope downward across the whole range of potential output, meaning that the most productively efficient outcome is for a single firm to serve the entire market.[139]

These industries are natural monopolies, in which competition is unworkable and undesirable, and regulation is necessary for efficiency. In a natural monopoly sector, neoclassical competition drives prices down to marginal cost, just as in other industries. But in natural monopoly sectors, unlike in other sectors, marginal cost pricing is by definition not enough to allow the firm to recover its average total costs. This means bankruptcy.[140] And even if multiple competing firms *could* be sustained (for example, through subsidies), that does not mean they *should* be. The downward-sloping marginal cost curve implies that productive efficiency is highest when just one firm serves the total market.[141] In these special cases, there is a justification for monopoly. There is also a justification for regulating that monopoly to ensure it does not charge inefficiently high prices.[142]

This neoclassical account of the economy suggested that the threat of industrial scale to classical liberalism might not be as severe as many feared. Decentralized competition among small businesses could continue to be the general principle of economic organization; monopolistic big business would be an exception. Antitrust law becomes the tool to break up monopoly and preserve competition in normal markets that can be served by multiple businesses without reducing productive efficiency. Public utility law

---

137. *Id.* at 265.

138. MARSHALL, *supra* note 126, at 265–66.

139. WARSH, *supra* note 64, at 77–129 (noting that monopoly is the logical consequence of unchecked increasing returns to scale and describing the features that Marshall included in his system to prevent that result from undermining competition).

140. ECONOMICS OF REGULATION, *supra* note 3, at 122; *see, e.g.*, *Natural Monopoly and Its Regulation*, *supra* note 1, at 548.

141. ECONOMICS OF REGULATION, *supra* note 3, at 74.

142. *See* MARSHALL, *supra* note 126, at 395; James Chen, *The Second Coming of Smyth v. Ames*, 77 TEX. L. REV. 1535, 1537 (1999); ECONOMICS OF REGULATION, *supra* note 3, at 74, 88, 106; FRANK, *supra* note 129, at 410–21.

becomes the tool to regulate the abnormal natural monopoly sectors, where increasing returns to scale make the productive efficiency of monopoly higher than that of competitive industries.

D. The Unpersuasiveness of the Natural Monopoly Synthesis

Given the enduring importance of the neoclassical concept of natural monopoly (a phenomenon I will describe below), it may be surprising that Marshall's contemporaries and immediate successors found his theory unpersuasive as a positive description of the economy. At best, Marshall's book described a transitional stage between Adam Smith's economy of individual proprietaries and a modern industrial economy. By the time neoclassical theory was widely known, big businesses benefiting from significant economies of scale were the general rule, not an exception.

Indeed, the "fixed cost controversy," also known as the "marginal cost controversy," shows the extent to which the neoclassical theory's flaws were apparent by the 1910s and 1920s.[143] Essentially, the controversy was what to do about the fact that most modern industries qualified as natural monopolies under the logic described above.[144] Marshall's elegant marginal cost pricing model only worked in industries where fixed costs were zero, but industrial production involves a relatively high ratio of fixed costs (machines, factories, fixed labor forces) to marginal costs (inputs to production, flexible labor).[145] In these sectors, competition will still drive prices down to short-run marginal cost, but marginal cost pricing will be too low to keep any firm with significant fixed costs in business. For some, this implied that we should treat every large industrial enterprise as a natural monopoly and impose public utility regulation throughout the economy.[146] Others proposed to "save" competition by using the government to subsidize fixed costs, thereby enabling multiple competitive firms to sustainably price at the marginal cost level.[147] Still others observed that the taxation required for subsidies was itself as inefficient as the deviation from marginal cost pricing, and also that

143. ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 308–322; Brett Frischmann & Christiaan Hogendorn, *Retrospectives: The Marginal Cost Controversy*, 29 J. ECON. PERSP. 193, 193 (2015).

144. ENTERPRISE AND AMERICAN LAW, *supra* note 39, at 308–09.

145. *Id.* Marshall himself dodged this issue in part with his famous "biological theory" of the firm, in which even firms with high economies of scale decline lose their advantage as they age, and are therefore subject to competition. MARSHALL, *supra* note 126, at 263.

146. CONTROL OF TRUSTS 1914, *supra* note 80, at 141 (describing the position held by many "business men [sic]").

147. *Id.* at 196 (describing Harold Hotelling's view).

such a system would result in an overproduction of high fixed cost goods relative to substitutes.[148]

In this way, the reality of early 20th century industrialization frustrated the neoclassical hope that classical competition might not need saving. Marshall succeeded in reinventing the discipline of economic science but failed to put the genie of industrial scale back in the bottle.

E. The Imperfect Competition Revolution

Edward Chamberlin's model of "monopolistic competition" and Joan Robinson's model of "imperfect competition," published contemporaneously in the 1930s, were the theoretical breakthrough that finally dissolved the problem of industrial scale.[149] These models posit that firms offer partially differentiated products that appeal to different tastes.[150] Firms enjoy a sort of monopoly power with respect to the market for their product.[151] They therefore can command prices that exceed marginal cost, saving them from the economic ruin that neoclassical models of competition predicted would befall firms suffering from "the fixed cost" problem.[152] At the same time, these firms do not completely foreclose competition in the manner of a neoclassical monopolist because their market power is limited by the existence of relatively close substitutes.[153] In short, where neoclassical theory sought to divide firms into competitive and monopolistic categories, imperfect competition theory regards almost all firms as part monopoly and part competitive.[154]

Chamberlin and Robinson opened the door to microeconomics without perfect competition at its center. In a world of imperfect competition, unlike in Marshall's elegant universe, there are tradeoffs. Neither the allocative efficiency promised by perfect competition nor the productive efficiency promised by monopoly is typically realized in full. And there is no *a priori* basis for supposing that a particular level of concentration is optimal in a particular sector. This was a revolution whose effect, in the words of economic historian Mark Blaug:

---

148. *Id.* at 197–99 (describing Ronald Coase's view).

149. ECONOMIC THEORY, *supra* note 57, at 375–79; WARSH, *supra* note 64, at 109, 111–15. J.M. Clark's efforts to find a "middle way" between the extremes of the public utilities vision and the antitrust vision, though written in a different style, were also an important predecessor to these ideas. CONTROL OF TRUSTS 1914, *supra* note 80, at vi (advocating a "third course" of "regulating competition").

150. ECONOMIC THEORY, *supra* note 57, at 375–79.

151. *Id.*

152. *Id.*

153. *Id.*

154. *Id.*

> [Was] to multiply the number of market structures that economics must analyse . . . . Policy prescriptions in a world of monopolistic competition and oligopoly cannot be based merely on the degree to which a particular market structure departs from the norms of perfect competition. Price theory has ever since been more complicated and less satisfying, and it is hardly surprising that some critics should not complain that we are left with little more than ad hoc theorizing. We can never go back to the bold generalities of Marshallian price theory.[155]

Into the vacuum left by the explosion of the Marshallian generalities stepped starkly divergent perspectives about what to do next. From one perspective, the existence of imperfect competition suggested a justification for aggressive government intervention throughout many areas of the economy.[156] Some midcentury economists, for example, favored treating advertising and aesthetic innovation (such as fins on automobiles) as anticompetitive practices that should give rise to antitrust liability. They thought that such practices "artificially" differentiate products that could otherwise be commodities produced under conditions more closely approximating perfect competition.[157] Joan Robinson herself drew even more radical conclusions: she saw her model as another nail in the coffin of *laissez-faire*, and a mandate for the radical reform of society.[158] After publishing her imperfect competition model, she turned to the study of Marx.[159]

But from another point of view—that of economists like Edward Chamberlin, John Maurice Clark, and Joseph Schumpeter—imperfect competition was not a failure to be corrected, but a fact about the world to be accepted, and not necessarily an unhappy one.[160] The new model of competition was only "imperfect" relative to an alternative called "perfect competition," but that alternative was illusory and would be productively inferior even if it existed. As Schumpeter put it in 1942: "[n]either Marshall . . . nor the classics saw that perfect competition is the exception

---

155. *Id.* at 379.

156. *Framing Chicago*, *supra* note 8, at 1878 (citing Joseph A. Schumpeter & A.J. Nichol, *Robinson's Economics of Imperfect Competition*, 42 J. POL. ECON. 249, 250–51 (1934)).

157. Herbert Hovenkamp, *United States Competition Policy in Crisis: 1890–1955*, 94 MINN. L. REV. 311, 339 (2009) [hereinafter *Competition Policy in Crisis*].

158. ZACHARY D. CARTER, THE PRICE OF PEACE: MONEY, DEMOCRACY, AND THE LIFE OF JOHN MAYNARD KEYNES 248–50, 414–16, 454–56 (2020).

159. JOAN ROBINSON, AN ESSAY ON MARXIAN ECONOMICS (2d ed. 1966).

160. Don Bellante, *Edward Chamberlin: Monopolistic Competition and Pareto Optimality*, 2 J. BUS. & ECON. RSCH. 17, 17 (2004).

and . . . even if it were the rule there would be much less reason for congratulation than one might think."[161]

Schumpeter's own contributions struck a different but related blow at neoclassical competition. Neoclassical models propose to describe how supply and demand determine prices in a static equilibrium, but Schumpeter recognized that "equilibrium" was just a hypothetical thought experiment, and perhaps not a particularly helpful one.[162] The most important kind of competition is not competition *in* a market, but competition *for* new markets—the rivalrous, restless innovation that disrupts equilibrium and drives economic growth.[163] Schumpeter thought that the monopolies (temporarily) captured by entrepreneurial firms had always been a feature of the capitalist system, and had become even more important in the industrial age.[164] The attempt to use government policy to enforce or approximate perfect competition was a fool's errand, for "firm[s] of the type that [are] compatible with perfect competition" were "in a less favorable position to evolve and to judge new possibilities" than big businesses.[165]

In the postwar era, this tolerant approach to big business proved a good fit for both the evidence and the public mood. Our worst fears about the inexorability of monopoly and all that it entailed for social and political organization had not been realized. Instead, many sectors of the U.S. economy settled into oligopolies similar to the type described by Chamberlin: Coke and Pepsi, UPS and FedEx, Boeing and Airbus, Visa and Mastercard, and so on. We came to see that competition of a certain kind would persist, and we were not on the road to serfdom. As the old Brandeisian economy receded further into historical memory, popular anxiety about big business evaporated.[166]

This revolution had strong implications for the balance of power between antitrust and public utility regulation. In a world of imperfect competition, antitrust persists, albeit as a technocratic tool for fine-tuning economic outcomes, not a revolutionary program for the revitalization of American democracy. Natural monopoly fares less well. It is marooned as a

---

161. SCHUMPETER, *supra* note 70, at 78.

162. *Id.*

163. *Id.* at 85.

164. *Id.*

165. *Id.* Schumpeter worked from different antecedents and in a different style from Chamberlin and Robinson, but decades later, Paul Romer's theoretical work on endogenous growth models would hint at the deep links between dynamic competition and imperfect competition. Charles I. Jones & Paul M. Romer, *The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital*, 2 AM. ECON. J.: MACROECONOMICS 224, 228 (2010) (explaining that endogenous growth models only work under conditions of imperfect competition); WARSH, s*upra* note 64, at 215 (providing a historical account of Romer's insight on the connection between growth and imperfect competition).

166. *See* HOFSTADTER, *supra* note 53, at 313.

mostly moot exception to a general rule (perfect competition) in which we no longer believe. Where neoclassical theory conceptualized natural monopoly as one of two industry structures—competition or monopoly—imperfect competition theory understands it instead as merely an endpoint on a long continuum of structures, running from perfect competition at one end through monopolistic competition and oligopoly to natural monopoly at the other extreme.[167] Even in the rare case in which we find an industry that we think falls bang on that natural monopoly endpoint, the difference in predicted performance between that firm and the "naturally oligopolistic" firms adjacent to it on the continuum is one of degree, not of category. Accordingly, there is little reason to single out a few "natural monopolies" for radically different treatment from the rest of the economy. As John Maurice Clark argued in an influential 1940 paper, we are better off pursuing "workable competition" in all sectors, giving antitrust the jurisdiction to police competitive imperfections as best as it can throughout the whole economy, and perhaps even to tolerate problematic market power in some industries for some time, trusting dynamic competition to erode it.[168]

The Harvard School of Industrial Organization, which dominated antitrust commentary in the middle of the 20th century, worked from this perspective.[169] For the Harvard School, the pursuit of workable competition meant abandoning deductive, categorical, neoclassical analysis in favor of pragmatically weighing the costs and benefits of competition and monopoly, both acknowledged to be imperfect.[170] It thus grappled with a host of diverse empirical phenomena that had been largely abstracted away by the deductive neoclassical approach. For example, in one of the era's leading texts, Joe Bain defined more than 15 distinct types of market structure, arrayed between the old poles of natural monopoly and perfect competition.[171] The Harvard School manifested a clear preference for antitrust as the regulator of most or all these structures. Bain relegated discussion of public utility regulation to the last few pages of his book and made it clear that he thought the contemporary extent of public utility regulation was riddled with

---

167. JOE S. BAIN, INDUSTRIAL ORGANIZATION 34 (1959); *accord* Joseph E. Stiglitz, Towards a More General Theory of Monopolistic Competition 3–4 (Oct. 1984) (unpublished manuscript) (available online from Princeton University) ("1. Most firms . . . have some degree of monopoly power. . . . 2. In most industries, there is not a natural monopoly . . . . Firms are thus embedded in a competitive environment.").

168. J. M. Clark, *Towards a Concept of Workable Competition,* 30 AM. ECON. REV. 241 (1940); *Competition Policy in Crisis*, *supra* note 157, at 321–22; Bellante, *supra* note 160, at 20.

169. *Framing Chicago*, *supra* note 8, at 1855.

170. *Id.* at 1854.

171. BAIN, *supra* note 167, at 34.

inconsistencies and deficiencies.[172] He proposed that "public utility regulation should not be extended to presently unregulated industries unless there is a very clear and conclusive demonstration that other regulatory devices [for example, antitrust] will not suffice to preserve or institute a reasonably workable competition in these unregulated industries."[173]

The old legal theory of public utility regulation suffered a strikingly parallel fate. When the Supreme Court finally yielded to the New Deal political order in the late 1930s,[174] its retreat from the general rule of *Lochner* mooted *Munn*'s exception for enterprises "affected by a public interest,"[175] just as economists' retreat from the general rule of perfect competition mooted the analogous natural monopoly exception. The path of this retreat was laid out in the famous dissents of Brandeis[176], Stone[177], and especially Holmes[178] from the *Munn* line of public utility cases, and adopted as law by *Nebbia* (1934),[179] *West Coast Hotel* (1937),[180] and *Hope Natural Gas* (1944).[181] With these decisions, our law ceased to recognize a "closed class or category of business affected with a public interest."[182] In the eyes of the Constitution, as in the eyes of economic theory, all businesses were henceforth part of one category.

---

172. *Id.* at 543–46.

173. *Id.* at 629.

174. In the first decades of the 20th century, the Supreme Court was a stalwart defender of the classical economic and social order. As one historian put it, "[s]ubstantive due process embalmed classical liberal economics in the appellate reports . . . and thus extended its influence preternaturally" for years after its flaws had become obvious to economists. WIECEK, *supra* note 73, at 135.

175. *Munn*, 94 U.S. at 126 (quoting Matthew Hale, *De Portibus Maris*, *in* A COLLECTION OF TRACTS RELATIVE TO THE LAW OF ENGLAND 45, 78 (Francis Hargrave ed., 1787)).

176. New State Ice Co. v. Liebmann, 285 U.S. 262, 302–03 (Brandeis, J., dissenting) ("In my opinion, the true principle is that the state's power extends to every regulation of any business reasonably required and appropriate for the public protection. I find in the due process clause no other limitation upon the character or the scope of regulation permissible.").

177. Tyson & Bro. v. Banton, 273 U.S. 418, 451 (1927) (Stone, J., dissenting) ("[B]usiness affected with a public interest . . . seems to me to be too vague and illusory to carry us very far on the way to a solution. It tends in use to become only a convenient expression for describing those businesses, regulation of which has been permitted in the past.").

178. *Id.* at 446 (Holmes, J., dissenting) ("[T]he proper course is to recognize that a state legislature can do whatever it sees fit to do unless it is restrained by some express prohibition in the Constitution . . . . [T]he notion that a business is clothed with a public interest and has been devoted to the public use is little more than a fiction intended to beautify what is disagreeable to the sufferers.").

179. Nebbia v. New York, 291 U.S. 502, 536 (1934) (eliminating the "constitutionalized" version of public utilities law by allowing economic regulation of businesses without a finding that they were affected by a public interest).

180. West Coast Hotel Co. v. Parrish, 300 U.S. 379, 400 (1937) (eliminating the practice of substantive due process review itself, to which *Munn* had been an exception, thereby doing away with the need for a public interest test).

181. Fed. Power Comm'n v. Hope Nat. Gas Co., 320 U.S. 591, 603 (1944).

182. *Nebbia*, 291 U.S. at 536.

II. THE STRANGE AND UNHELPFUL AFTERLIFE OF THE NATURAL MONOPOLY CONCEPT, 1950–2025

As of the middle of the 20th century, an observer who had imbibed the insights of the imperfect competition revolution would have been justified in thinking the idea of natural monopoly as traditionally understood was obsolete. Such an observer might have predicted that the concept should either disappear entirely along with the concept of perfect competition to which it was an exception, or be reimagined to better describe the relationship between natural monopoly and the spectrum of imperfectly competitive industry structures.

What happened instead was more nuanced and surprising. We *did* stop applying public utility regulation to new technologies, just as Joe Bain suggested.[183] In fact, we went further: between the 1970s and 1990s, we deregulated many of the industries that we had previously treated as public utilities.[184] But the concept of natural monopoly did not disappear. Nor was it updated to better describe a world of imperfect competition. Instead, for reasons mostly internal to the fascinations of economic science, we came to use a neoclassical vocabulary—including, sometimes, the natural monopoly concept—to describe an imperfectly competitive reality.[185] Thus, a central concern of the deregulatory era was to separate assets that were not natural monopolies (and therefore should be deregulated) from assets that were natural monopolies (and therefore should arguably remain regulated).[186] Unfortunately, our neoclassical natural monopoly concept offered little help in drawing the line between natural monopoly and imperfect competition, which contributed to several deregulatory false starts and costly disasters.[187] Today, the neoclassical notion of natural monopoly remains "on the books," but almost no one finds it useful for diagnosing or remedying the industrial organization problems of our own time.[188] In this Part II, I recount the intellectual history of how our ideas about natural monopoly took on this strange and unhelpful form.

183. *See* BAIN, *supra* note 167, at 629.
184. *Infra* Part II.B.
185. *Infra* Part II.A.
186. *Infra* Part II.B.
187. *Infra* Part II.B.
188. *Infra* Part II.C.

A. Neoclassical Vocabulary for an Imperfectly Competitive Reality

The Chicago School was famously founded on hostility to the imperfect competition revolution and is sometimes credited with helping to drive a return to neoclassical models of competition.[189] For the purposes of this Article, it is important to understand exactly in what sense this is true. Chicagoans' antipathy to imperfect competition was fundamentally methodological and rhetorical: they accepted the phenomenon as a fact about the world and derived from that fact the same doubts as the Harvard School about the value of public utility regulation (plus additional, stronger ones). They merely thought the neoclassical system was a better vocabulary and methodology for economic analysis.

This core perspective was evident from the beginning of the movement. In a foundational 1953 paper, Milton Friedman argued that economic models should be judged by their ability to make accurate predictions about economic phenomena (such as prices and interest rates), not by the realism of their assumptions as a description of the world.[190] Friedman thought that imperfect competition models inappropriately privileged realism over predictive power.[191] Imperfect competition theory "was explicitly motivated, and its wide acceptance and approval largely explained, by the belief that the assumptions of 'perfect competition' or 'perfect monopoly' said to underlie neoclassical economic theory are a false image of reality," not by any shortcomings in the predictive power of neoclassical models.[192] According to Friedman, those who wanted good predictions about economic phenomena would be better served by the neoclassical price theory set out in Marshall's 1890 *Principles of Economics*, which "seems to me both extremely fruitful and deserving of much confidence for the kind of economic system that characterizes Western nations."[193] In other words, as we use Newtonian physics to design buildings even after it has been (in some sense) "disproven" by Einstein, so we can use Marshallian perfect competition even after Chamberlin.[194]

The Chicago School's response to imperfect competition—accept it in fact while downplaying it in rhetoric—was in many respects merely a more forthright and pugnacious version of the attitude taken by the 20th-century economics profession at large. This is evident in Paul Samuelson's 1948

189. *Framing Chicago*, *supra* note 8, at 1846–47.
190. MILTON FRIEDMAN, ESSAYS IN POSITIVE ECONOMICS 15 (1953).
191. *Id.* at 15.
192. *Id.*
193. *Id.* at 41–42.
194. *See id.* at 38.

textbook, *Economics,* which came to define the mathematical, center-left mainstream of the discipline. Samuelson acknowledged that economics had gone through an imperfect revolution in the 1930s and disagreed with Friedman's criticisms.[195] Yet Samuelson did not go so far as to abandon perfect competition in favor of monopolistic competition. Instead, his textbook hewed to the same pattern as John Stuart Mill's classical and Marshall's neoclassical textbooks: it put perfect competition and its advantages at the center of microeconomic thought, treating imperfect competition and similar phenomena as "market failures" that deviated from this ideal, but about which there was not a lot to say.[196] As a description of what academic economists did, this was not wrong: much of the progress of economic science during the 20th century continued to elaborate the theory of perfect competition, now treated more as a thought experiment than an accurate description of reality. For example, among the most cherished triumphs of the 20th century was Kenneth Arrow and Gerard Debreu's proof: under assumptions of perfect competition, every competitive equilibrium is "Pareto-efficient," thereby bringing mathematical rigor to the "invisible hand theorem" that had been at the center of both classical and neoclassical economics.[197]

This methodological resurrection of neoclassical thought gave Chicago School thinkers like Robert Bork and Richard Posner a vantage from which to portray the Harvard School antitrust commentary (mostly unfairly) as economically unscientific and (somewhat fairly) ad hoc.[198] Compared to the Harvard School style, the Chicago School's work was simpler, more deductive, and more deterministic.[199] It offered clear recommendations for

---

195. Paul A. Samuelson, *The Monopolistic Competition Revolution*, *in* 3 THE COLLECTED SCIENTIFIC PAPERS OF PAUL A. SAMUELSON 18, 37–51 (Robert C. Merton ed., 1972).

196. PAUL SAMUELSON, ECONOMICS: AN INTRODUCTORY ANALYSIS 445–51, 491–507 (1948); *see* ZACHARY D. CARTER, THE PRICE OF PEACE: MONEY, DEMOCRACY, AND THE LIFE OF JOHN MAYNARD KEYNES 378, 415–16 (2020); James K. Galbraith, *Keynes, Einstein, and Scientific Revolution*, AM. PROSPECT, Winter 1994, at 62, 66 ("What Samuelson did . . . was to push the daemon of Keynesian relativity back into its box.").

197. *Fundamental Theorems*, *supra* note 131, at 200. Pareto efficiency is itself a sort of downgrade from earlier efforts to prove the efficiency of competition. A Pareto efficient outcome is merely the best outcome that can be achieved *by voluntary transactions*, which makes it a somewhat tautological metric for assessing the optimality of a system of voluntary market exchange. *See* DANIEL M. HAUSMAN & MICHAEL S. MCPHERSON, ECONOMIC ANALYSIS, MORAL PHILOSOPHY, AND PUBLIC POLICY 87–90 (1996).

198. Herbert Hovenkamp, *Antitrust Policy After Chicago*, 84 MICH. L. REV. 213, 217–18 (1985) [hereinafter *Antitrust Policy After Chicago*]; *see, e.g.*, Richard A. Posner, *The Chicago School of Antitrust*, 127 U. PA. L. REV. 925, 932–33 (1979).

199. *Antitrust Policy After Chicago*, *supra* note 198, at 225; ROBERT H. BORK, THE ANTITRUST PARADOX, at xiii (2d ed. 1993) ("Price theory is not only a powerful tool of analysis, it is also a powerful

the conduct that justified antitrust intervention (cartels, "naked" agreements in restraint of trade, and horizontal mergers between businesses with very large market shares) from the conduct that didn't (most of the vertical mergers, agreements, and practices that antitrust had previously viewed with suspicion).[200]

Bork and Posner attributed their positions to the application of Marshallian price theory,[201] but much of what is novel in their work is actually derived from their acceptance of the superior productive efficiency of big business—a position that Marshall himself had denied and that the Harvard School had accepted.[202] For example, Bork and Posner argued that monopolies and oligopolies achieved through internal growth should be tolerated because such growth was likely caused by productive efficiencies superior to their competitors.[203] Purely vertical mergers should be tolerated because they are likely inspired by the desire to increase productive efficiency, rather than the desire for market power.[204] The same possibility of productive efficiency through vertical integration justifies "vertical price fixing (resale price maintenance), vertical market division (closed dealer territories), and, indeed, all vertical restraints," including tying, which "are beneficial to consumers and should for that reason be completely lawful."[205] Even price fixing and market division arrangements may be tolerated if they are ancillary to a cooperative arrangement plausibly capable of creating productive efficiency.[206]

Frank Easterbrook was more candid than his Chicago School colleagues that the school worked within the paradigm of accommodationism to an imperfectly competitive world.[207] Referencing John Maurice Clark, Easterbrook thought the Chicago School should be called the "Workable Antitrust Policy School," insofar as its main program was to cleanse antitrust of the outdated tendency to "condemn every practice that did not look like hearty yeomen competing from moment to moment" and of its continued fealty to "the model of atomistic competition."[208] Easterbrook held that the Chicago School's first fundamental insight was that "[n]o antitrust policy

---

form of rhetoric. . . . [T]he simplest ideas are also the most powerful and entirely adequate to the tasks of the law. . . . It was possible to win arguments and do so decisively.").

200. RICHARD A. POSNER, ANTITRUST LAW 52, 132, 224, 195, 236, 242 (2d ed. 2002).

201. BORK, *supra* note 199, at 67–68; POSNER, *supra* note 200, at 932.

202. BAIN, *supra* note 167, at 155.

203. *Id.* at 196–97.

204. *Id.* at 245.

205. BORK, *supra* note 199, at 297–98.

206. *Id* at 279.

207. Frank H. Easterbrook, *Workable Antitrust Policy*, 84 MICH. L. REV. 1696, 1702–06 (1986).

208. *Id.* at 1707.

should be based on a belief that atomistic competition is better than some blend of cooperation and competition" and that "[t]he right blend varies from market to market"[209]—a proposition Joe Bain might have agreed with.

Richard A. Posner's 1968 article, *Natural Monopoly and Its Regulation*—still the most thorough legal academic article about natural monopoly ever written—was similarly alert to the implications of imperfect competition.[210] Posner defined his subject in simple neoclassical terms: "[i]f the entire demand within a relevant market can be satisfied at lowest cost by one firm rather than by two or more, the market is a natural monopoly . . . ."[211] His analysis of what should be done about natural monopoly, however, acknowledged and engaged the much more nuanced reality of imperfect competition. Posner observed that there is less to fear from monopoly than neoclassical models previously predicted because the real-world alternative to monopoly pricing in our economy is generally oligopolistic pricing, not perfect competition pricing.[212] Moreover, fear of dynamic competition from entrepreneurs who "devise ingenious methods of challenging or supplanting the monopolist" may limit the monopoly's ability to charge prices significantly in excess of the oligopolistic level.[213] Posner proposed on that basis, like Schumpeter before him, that we simply tolerate the imperfections of the competitive process, even if they lead naturally, in some cases for some periods of time, to monopoly.[214] Posner acknowledged, however, that this outcome was politically unlikely.[215] As a second choice, therefore, he advocated the "somewhat more realistic objective" of "deregulation of those industries that are not natural monopolies, such as natural gas production, aviation, and trucking."[216]

Therefore, if we abstract away the Chicago School's methodological and rhetorical preferences, what remains is a perspective on natural monopoly that is fundamentally consistent with the Harvard School's: natural monopoly regulation is often misguided and should be disfavored relative to antitrust enforcement. Chicago's innovation was to show us how to describe imperfectly competitive reality in a neoclassical vocabulary, and natural monopoly was a feature of that vocabulary—even if Chicagoans deployed

---

209. *Id.* at 1700 (noting that by "cooperation," Easterbrook meant something like "vertical integration").

210. *Natural Monopoly and Its Regulation*, *supra* note 1, at 573; *see* MILTON FRIEDMAN, CAPITALISM AND FREEDOM 128 (1962) [hereinafter CAPITALISM AND FREEDOM].

211. *Natural Monopoly and Its Regulation*, *supra* note 1, at 548.

212. *Id.* at 560.

213. *Id.* at 558.

214. *Id.* at 561.

215. *Id.* at 638–39.

216. *Id.* at 639.

the concept mostly to argue that a particular industry didn't quality as a natural monopoly, or that its regulation was unnecessary.

B. The Deregulatory Era

Between the 1970s and the 1990s, U.S. policymakers implemented Posner's second choice idea: prune away regulation from functions that were no longer perceived to be natural monopolies while continuing to apply it (in some form) to functions that *were* still thought to be natural monopolies.[217] This was doubly ironic. One irony was that the motivation for deregulation was born out of the insights of the imperfect competition revolution, but the movement's scope and tactics were guided by the thoroughly neoclassical concept of natural monopoly. Another irony was that the deregulatory era became the heyday of natural monopoly as a practical guide to regulation. The public utility regulatory edifice was assembled by lawyers before the concept of natural monopoly was widely understood; now, at least, the concept was available to attend to the edifice's disassembly.

Deregulation had both successes and failures. Transportation deregulation was a success story. By the 1970s, most observers had concluded that airlines, trucking, and railroads were no longer natural monopolies, in part because of the possibility of "multimodal" competition amongst the various sectors.[218] Regulators thus extricated themselves from the business of setting cost-of-service rates in these sectors, allowing competition to establish prices instead.[219] The airline's price regulator (the Civil Aeronautics Board) was shuttered entirely.[220] The results were striking: in the air travel sector, the U.S. Government Accountability Office estimated a 40% reduction in prices between 1980 and 2006, and a corresponding expansion in traffic.[221]

---

217. Kearney & Merrill, *supra* note 4, at 1323, 1328–29 (describing the replacement of a paradigm focused on regulatory oversight of particular industries irrespective of "whether the regulated industry was naturally competitive or a natural monopoly" with a new paradigm focused on regulating "bottleneck' monopolies"). Some of the early rhetoric that surrounded the envisioned movement was more nakedly deregulatory. *See, e.g.*, Ronald Reagan, First Inaugural Address (Jan. 20, 1981) ("In this present crisis, government is not the solution to our problem; government is the problem."); CAPITALISM AND FREEDOM, *supra* note 210, at 128–29 (making an influential case for deregulation). But as it grappled with the complex technical realities of the regulated industries, the movement morphed into something more nuanced.

218. *See, e.g.*, *Natural Monopoly and Its Regulation, supra* note 1, at 639.

219. Kearney & Merrill, *supra* note 4, at 1336–37.

220. *Id.* at 1335.

221. U.S. GOV'T ACCOUNTABILITY OFF., GAO-06-630, AIRLINE DEREGULATION, REGULATING THE AIRLINE INDUSTRY WOULD LIKELY REVERSE CONSUMER BENEFITS AND NOT SAVE AIRLINE PENSIONS 19 (2006). For a more nuanced analysis of the success of airline deregulation, *see* Severin Borenstein & Nancy L. Rose, *How Airline Markets Work . . . or Do They? Regulatory Reform in*

The earliest phase of telecommunications deregulation was also a success story. Reformers recognized local telephone poles and wires infrastructure as a natural monopoly but postulated that other functions that had traditionally been vertically integrated into the great AT&T monopoly were potentially competitive, including long-distance lines and telephone equipment. After more than a decade of resisting efforts by the Federal Communications Commission (FCC) and Department of Justice (DOJ) to force competition into these markets,[222] AT&T entered into a consent decree in 1982 with the DOJ under which it spun off its local phone networks into seven regional "Baby Bells," which were forced to interconnect not only with AT&T's remaining long-distance business, but also with new long-distance competitors.[223] For some observers, this approach announced a "Bell Doctrine" or "Baxter Doctrine," which held that regulators should "'quarantine' the regulated monopoly segment of the industry by separating its ownership and control from that of the firms in potentially competitive segments of the industry."[224] Competition in the long distance market drove significant cost decreases in long-distance telephone service.[225] More indirectly, the consent decree is often credited with replacing AT&T's sclerotic monopoly with the Schumpeterian competitive landscape that gave us cost-effective cellular phone service, and perhaps even the rise of widespread internet service.[226]

Subsequent attempts to bring competition to local telephone service, however, were less successful. These further reforms were initiated by the FCC and state public utility commission,[227] and culminated in the 1996 Telecommunications Act. The Act mandated that the Baby Bells offer competing local phone service retailers access to their physical infrastructure.[228] The idea was to allow new entrants to essentially "resell"

---

the Airline Industry, *in* ECONOMIC REGULATION AND ITS REFORM: WHAT HAVE WE LEARNED? 63, 129 (Nancy L. Rose ed., 2014).

222. TIM WU, THE MASTER SWITCH: THE RISE AND FALL OF INFORMATION EMPIRES 202–03 (2010) [hereinafter THE MASTER SWITCH].

223. Howard A. Shelanski, *Adjusting Regulation to Competition: Toward a New Model for U.S. Telecommunications Policy*, 24 YALE J. ON REG. 55, 62 (2007).

224. Paul L. Joskow & Roger G. Noll, *The Bell Doctrine: Applications in Telecommunications, Electricity, and Other Network Industries*, 51 STAN. L. REV. 1249, 1249–50 (1999).

225. Robert W. Crandall, *The Failure of Structural Remedies in Sherman Act Monopolization Cases* 70, 71, 83 (AEI-Brookings Joint Ctr., Working Paper No. 01-05, 2001) (arguing that restructuring of the telephone system reduced rates, though similar reductions could have been achieved through an interoperability rule without the breakup of AT&T).

226. THE MASTER SWITCH, *supra* note 222, at 54, 190.

227. *Id.* at 110, 194 (including the deregulation of wireless service).

228. *Id.* at 194; WILLIAM J. BAUMOL & J. GREGORY SIDAK, TOWARD COMPETITION IN LOCAL TELEPHONY 7, 9 (1994); James B. Speta, *Deregulating Telecommunications in Internet Time*, 61 WASH.

local phone service delivered by the incumbents' infrastructure.[229] However, this idea did not seem to generate many practical benefits.[230] Its most famous outcome was a great deal of complex, acrimonious regulatory proceedings.[231] Unbundling required regulators to make controversial decisions about how the unbundled services should be priced.[232] The vertically integrated incumbents had a strong incentive to overprice the services in order to preserve their monopoly against competition.[233] But if regulators set prices of unbundled services too low, the new entrants would free-ride on the incumbent's backbone investments, leading to an unsustainable system.[234]

Electricity deregulation was even more complicated and less successful. Traditionally, vertically integrated electric utilities generated, distributed, and sold electricity.[235] Analogies to telecommunications restructuring now suggest that electricity distribution over poles and wires might be the only true natural monopoly function, with generation and retailing seen as potentially competitive.[236] In the 1990s, the Federal Energy Regulatory Commission (FERC) pushed to restructure the industry along these lines, though for complicated reasons, FERC could only implement some aspects of this vision while others required state action.[237] When the dust settled, California, Texas, and states in the Northeast and Midwest had substantially restructured their systems.[238] The rest of the country mostly stuck to the traditional, vertically integrated, regulated utility model.

The earliest attempts to duplicate the AT&T breakup sought to enable "wheeling" of electricity from competitive generators to competitive users

---

& LEE L. REV. 1063, 1094–95 (2004) (explaining the influence of Baumol & Sidak's position on the shape of the 1996 Telecommunications Act).

229. BAUMOL & SIDAK, *supra* note 228, at 12.

230. Jerry Hausman & Gregory Sidak, *Did Mandatory Unbundling Achieve its Purpose? Empirical Evidence from Five Countries* 70–71 (MIT Dep't of Econs., Working Paper No. 04-40, 2004).

231. Yoo, *supra* note 4, at 857.

232. *See* ALFRED KAHN, WHOM THE GODS WOULD DESTROY, OR HOW NOT TO DEREGULATE 3–11 (2001) [hereinafter HOW NOT TO DEREGULATE] (describing complex controversies in the unbundling of telephone service); Tyler McNish, *Reform Incentives, Transform the Grid: Making Good on Hawai`i's Renewable Energy Ambitions*, 45 ECOLOGY L.Q. 583, 628 (2019).

233. HOW NOT TO DEREGULATE, *supra* note 232, at 17.

234. *See id.* at 15.

235. McNish, *supra* note 232, at 624–29 (2019).

236. Kearney & Merrill, *Great Transformation*, *supra* note 4, at 1385 (describing the idea of electricity restructuring and noting that "[i]t is too early to assess the efficiency effects of these reforms, but the expectation is that lower average prices and net welfare gains will be substantial"); THE END OF A NATURAL MONOPOLY: DEREGULATION AND COMPETITION IN THE ELECTRIC POWER INDUSTRY 1–3 (Peter Z. Grossman & Daniel H. Cole eds., 2003); SALLY HUNT, MAKING COMPETITION WORK IN ELECTRICITY 2, 3, 6 (2002).

237. Kearny & Merrill, *supra* note 4, at 1354.

238. Severin Borenstein & James Bushnell, *The U.S. Electricity Industry after 20 Years of Restructuring* 7 (MIT Dep't of Econs., Working Paper No. 252, 2015).

via a natural monopoly "poles and wires" utility.[239] These attempts generated acrimonious access pricing disputes similar to the local telephone unbundling controversies.[240] Even worse, the wheeling concept turned out to be fundamentally at odds with the fact that the electricity network did not work like a switched telecommunications network.[241] The better analogy for electricity is hydraulic: the electricity pumped into the overall system needs to match electricity outflows on an instant-by-instant basis; otherwise, the mismatch creates voltage and frequency excursions with devastating consequences for the overall system.[242] Vertically integrated electric utilities had traditionally solved this problem by planning and controlling a portfolio of vertically integrated power plants sufficient to cover all anticipated demand scenarios.[243] This function could not simply be turned over to a competitive market because markets do not instantaneously equilibrate supply and demand: price signals take time to do their work, and that work often requires painful bouts of gluts and scarcity.[244] Instead, the states that restructured were forced to acknowledge that system control was a natural monopoly and turn it over to a quasi-public independent system operator.[245]. The independent system operator might run a reverse auction system into which competitive generation companies sell their electricity, but it also retains the power to dispatch plants as needed to meet system demands.[246]

The result of this restructuring was a more complicated system than the traditional vertically integrated system that only partially resembled a competitive market. The benefits of this system are not obvious. The states that preserved the old, vertically integrated utility model have lower prices than the restructured states.[247] Moreover, the systems set up by the restructured states to balance electricity demand have failed twice—

239. Kearny & Merrill, *supra* note 4, at 1354.

240. McNish, *supra* note 232, at 628 (2019); Timothy P. Duane, *Regulation's Rationale: Learning from the California Energy Crisis*, 19 YALE J. ON REG. 471, 486 (2002); RICHARD F. HIRSH, POWER LOSS: THE ORIGINS OF DEREGULATION AND RESTRUCTURING IN THE AMERICAN ELECTRIC UTILITY SYSTEM 125–31 (1999); Richard D. Cudahy, *PURPA: The Intersection of Competition and Regulatory Policy*, 16 ENERGY L.J. 419, 423–24, 431–33 (1995).

241. HUNT, *supra* note 236, at V; WILLIAM HOGAN, MARKET DESIGN AND ELECTRICITY RESTRUCTURING 16 (2005).

242. HUNT, *supra* note 236, at 20.

243. *Id.* at 20, 24.

244. *See id.* at 33–34 ("Electricity really is different from other commodities in its need for short term coordination."); McNish, *supra* note 232, at 633 (2019).

245. HUNT, *supra* note 236, at 37.

246. HUNT, *supra* note 236, at 37–38, 41, 46, 58 (describing the "standard prescription" for ISO-based restructuring); Borenstein & Bushnell, *supra* note 238, at 4.

247. Borenstein & Bushnell, *supra* note 238, at 1, 17, 18.

California in 2001 and Texas in 2021—with high and avoidable political, social, and economic costs.[248]

Taken as a whole, the experience of the deregulatory era supports three propositions. First, it validated the general insight that imperfect competition *usually* outperforms regulated monopoly.[249] Airline service, trucking, and long-distance telephony, for example, perform better as imperfectly competitive markets than regulated utilities. Second, natural monopoly is a real phenomenon, even if it may be rarer and more difficult to regulate than we once thought. As reformers advanced from their early wins (transportation) to trickier attempts to unbundle local telephony and electricity distribution, they bumped into functions or assets that really did seem to work better as monopolies than as competitive markets[250] or that could not be efficiently unbundled from related non-natural monopoly assets.[251] Third, and most importantly, we are not particularly good at identifying these natural monopolies. In some industries (for example, transportation), we correctly concluded that no natural monopoly was at stake, but in others (for example, telecommunications and electricity), our hypotheses were not correct about which functions could be made competitive and which were natural monopolies.

C. The Contemporary Irrelevance of Natural Monopoly

The deregulatory era was the natural monopoly concept's last gasp of relevance. In the quarter century since, the concept's popularity and credibility have descended to a nadir. Natural monopoly remains "on the books" as a feature of our neoclassical economic vocabulary, and it is not

248. Severin Borenstein, *The Trouble with Electricity Markets: Understanding California's Restructuring Disaster*, 16 J. ECON. PERSP. 191, 191 (2002); Dean Jepsen, *Examining the 2021 Texas Power Grid Crisis*, 27 PUB. INT. L. REP. 23, 23–25 (2021).

249. *See, e.g.*, Howard A. Shelanski, *Adjusting Regulation to Competition: Toward a New Model for U.S. Telecommunications Policy*, 24 YALE J. ON REG. 55, 56 (2007) ("[E]x ante regulation that depends for its rationale on monopoly market structure should give way to ex post intervention against specific, anti-competitive acts on the model of conventional antitrust . . . . [H]owever . . . interconnection [regulation] still [has] a role to play in advancing telecommunications policy objectives."); Yoo, *supra* note 4, at 867 (articulating "a choice between two regulatory paradigms, one that focuses on breaking down the monopoly by stimulating competitive entry and another that surrenders to the monopoly and simply seeks to allocate [access to it]," and arguing that policymakers should choose the first course except in rare instances where competition is infeasible).

250. *See* HOW NOT TO DEREGULATE, *supra* note 232, at 1.

251. HUNT, *supra* note 236, at 25; Paul L. Joskow, *The Role of Transaction Cost Economics in Antitrust and Public Utility Regulatory Policies*, 7 J.L. ECON. & ORG. 53, 66–67 (1991); PAUL L. JOSKOW & RICHARD SCHMALENSEE, MARKETS FOR POWER: AN ANALYSIS OF ELECTRICAL UTILITY DEREGULATION 120 (1983); HOW NOT TO DEREGULATE, *supra* note 232, at 8.

uncommon for scholars to gesture at its existence in passing—but it almost never forms a part of their core analysis or program of reform.[252]

Indeed, the novel vocabularies that 21st-century scholars use to describe problems that might once have been seen through the lens of natural monopoly are a clue to just how irrelevant the concept of natural monopoly has become. In the controversy over what to do about potential discrimination by broadband internet service providers—the first major industrial organization problem of the internet era—no one wanted to talk about natural monopoly.[253] As Tim Wu pointed out, the main participants in this debate shared a common "Schumpeterian" perspective.[254] By that, he meant that neither side cared much for correcting imperfections in the mechanism of price competition (the traditional justification for natural monopoly).[255] Thinkers like Wu preferred some form of intervention, because they worried the cable companies that controlled the "middle" part of the internet's architecture might vertically integrate with the "ends" of that architecture, favoring and disfavoring content based on the cable companies' economic interests, thereby stifling dynamic innovation.[256] These interventionist thinkers, however, avoided the "natural monopoly" vocabulary that previous generations had used to explain similar bottleneck phenomena in the railroad, electricity, and telecommunications sectors in favor of industry-specific neologisms like "end to end," "network neutrality," and "generative."[257] Those who opposed intervention also found little use for natural monopoly—instead, they argued that the new internet sector, like the rest of our dynamically competitive economy, was best regulated by antitrust alone.[258] The debate between these two sides thus reflected a consensus judgment that whether or not broadband internet service was a natural monopoly was mostly beside the point.

When the plucky "edge providers" (such as Google and Facebook) that the net neutrality advocates had been concerned to protect in the 2000s grew

---

252. *Supra* notes 18–20.

253. Tim Wu, The Broadband Debate, A User's Guide, 3 J. On Telecomm. & High Tech. L. 69, 76 (2004) (mentioning "natural monopoly" only once in a comprehensive summary of the debate).

254. *Id.* at 70, 80.

255. *Id.*

256. *Id.* at 73–74, 79.

257. *See* Tim Wu & Christopher S. Yoo, *Keeping the Internet Neutral?: Tim Wu and Christopher Yoo Debate*, 59 FED. COMM. L.J. 575, 585 (2007) (stating that the application of natural monopoly is "a big debate that I don't want to get into"); Bruce M. Owen, *Antecedents to Net Neutrality*, 30 REGUL. 14, 15–16 (2007) (noting that net neutrality advocates "invent new language to describe both a familiar economic problem and a familiar legal and regulatory solution to that problem.").

258. Owen, *supra* note 257, at 17 (2007); Christopher S. Yoo, *Would Mandating Broadband Network Neutrality Help or Hurt Competition? A Comment on the End-to-End Debate*, 3 J. ON TELECOMM. & HIGH TECH. L. 23, 61–62 (2004).

into the sprawling "Big Tech" enterprises of the 2010s and 2020s, the main participants in the discourse on what to do—with the important exception of Herbert Hovenkamp[259]—have again found little use for the natural monopoly vocabulary. Instead, the novel coinages that earlier characterized the net-neutrality debate have continued to proliferate, with the interventionist vanguard speaking of "dominant digital platforms," "social infrastructures," "winner-take-all markets," "the separation of platforms and commerce," "material preferencing," "firewalling," and "the critical and competitive significance of data."[260] Some writers allude to the existence of the natural monopoly concept, but they tend to treat it as passé and almost never apply it to the facts at hand.[261]

The dominant trend is to see competition, not regulated monopoly, as the solution for problems of industrial organization. The influential New Brandeis movement, and the broader antimonopoly left of which it is part, exemplify this phenomenon. New Brandeisians advocate the revitalization of antitrust enforcement to break up monopoly and reinvigorate competition.[262] They appropriately reject the deregulatory bias that the Chicago School infused into contemporary antitrust doctrine,[263] yet follow Chicago in its use of an unrepentantly neoclassical economic vocabulary that romanticizes competition, and exceed Chicago in their insensitivity to the insights of the imperfect competition revolution.[264] Specifically, the movement's emphasis on what Brandeis called the "Curse of Bigness" leaves little room for the "Blessings of Bigness" perceived by thinkers like Edward Chamberlin, Joseph Schumpeter, John Maurice Clark, and Joe Bain. In a discourse so committed to the virtue of competition, it is difficult for the concept of natural monopoly—good monopoly—to gain any purchase.

Neglect of natural monopoly is not limited to the problems of internet regulation. Consider the burgeoning "grid governance" literature, which analyzes the potential of institutional reform to accelerate decarbonization of

---

259. *Antitrust and Platform Monopoly*, *supra* note 8, at 1971–72.

260. *Supra* notes 12–18.

261. *Amazon's Antitrust Paradox*, *supra* note 7, at 800 (arguing that Big Tech problems are not likely to be addressed by natural monopoly regulation because "critics challenged the theory of natural monopoly as an ongoing rationale for regulation"); K. Sabeel Rahman, *Regulating Informational Infrastructure: Internet Platforms as the New Public Utilities*, 2 GEO. L. TECH. REV. 234, 236 (2018) [hereinafter *Regulating Informational Infrastructure*] (arguing that "infrastructure can be conceptualized in much broader terms" than "natural monopoly").

262. *The New Brandeis Movement*, *supra* note 20.

263. *Id.*; Lina M. Khan, *The End of Antitrust History Revisited*, 133 HARV. L. REV. 1655, 1661–62, 1676–77 (2020); *Separation of Platforms and Commerce, supra* note 7, at 1022; *Antitrust's Unconventional Politics*, *supra* note 8, at 123; *Framing Chicago*, *supra* note 8, at 1878.

264. ANTITRUST IN THE NEW GILDED AGE, *supra* note 7, at 70.

the electricity grid.[265] Like the literature on Big Tech, this literature brims with industry-specific jargon and fact-intensive analyses. Perspectives on what is to be done range from incrementalist, market-oriented proposals[266] to more radical "rebuilding" proposals,[267] to proposals for the decommodification of electricity, possibly through revitalization of public utility regulation.[268] Absent from this debate, however, is the kind of work that a time traveler from the 20th century might have most expected: an effort to identify the natural monopoly features of the industry, and analyze whether monopoly might be delaying progress on decarbonization or increasing its cost.

Recent new cross-industry syntheses similarly downplay natural monopoly. Brett Frischmann's *Infrastructure*, for example, aspires to rebuild the rationale for our law's special treatment of public utility sectors on a "demand side" theory, in contrast to the "supply side" theory of natural monopoly.[269] Another new synthesis reimagines the field of public utility regulation as "Networks, Platforms, and Utilities."[270] This new field treats natural monopoly as a phenomenon that is continuous with "natural oligopoly," and is just one of six disjunctive factors that might qualify "NPU firms" for extraordinary legal treatment—a striking (though understandable) demotion for a concept that was once understood to be at the heart of the whole public utility regulatory field.

III. NATURAL MONOPOLY FOR A WORLD OF IMPERFECT COMPETITION

The contemporary irrelevance of natural monopoly puts us at a fork in the road. Should we disavow natural monopoly as a matter of theory, as we have already (mostly) abandoned it in practice? Or should we attempt to renovate the theory to make it more relevant to contemporary problems of industrial organization? In this Part III, I take the latter path. The deregulatory experience recounted above confirms that natural monopoly is a real phenomenon, even if it may be rarer, harder to identify, and more dangerous to regulate than we once thought. And if natural monopoly is a

265. Shelley Welton, *Rethinking Grid Governance for the Climate Change Era*, 109 CAL. L. REV. 209 (2021); Alexandra Klass et al., *Grid Reliability Through Clean Energy*, 74 STAN. L. REV. 969, 970 (2022); Joel B. Eisen & Heather E. Payne, *Rebuilding Grid Governance*, 48 BYU L. Rev. 1057 (2023); William Boyd, *Decommodifying Electricity*, 97 S. CAL. L. REV. 101 (2024).

266. Welton, *supra* note 265, at 210.

267. Eisen, *supra* note 265, at 1057.

268. Boyd, *supra* note 265, at 947–48.

269. FRISCHMANN, *supra* note 13, at 108; Lawarence Lessig, *Reply, Re-Marking the Progress in Frischmann*, 89 MINN. L. REV. 1031, 1036 (2005) (glossing Frischmann with a useful flow chart).

270. RICKS ET AL., *supra* note 15, at 8–9.

problem, antitrust alone won't solve it, for a natural monopoly won't need anticompetitive conduct (the *sine qua non* of antitrust liability) to secure and perpetuate its advantage. Thus, we need a better test for natural monopoly, one that is more appropriate for a world of imperfect competition. Below, I develop such an updated test for natural monopoly, which I hope will allow it to resume its proper role in guiding the application of public utility regulation to problematic technological assets.

A. Our Current Natural Monopoly Concept Is Too Neoclassical

To make natural monopoly theory more relevant, we must first diagnose the reason for its irrelevance. The main reason, I think, is that our understanding of natural monopoly remains "too neoclassical." Consider a few typical definitions of natural monopoly from recent legal academic work:

> [1] Natural monopoly occurs when a single firm can serve the entire market more cheaply than can two firms, a condition known as "subadditivity." A sufficient condition for subadditivity is the existence of scale economies throughout the entire range of production, such as occurs when fixed costs are very high.[271]

> [2] The structural circumstances in which monopoly is the cheapest way of organizing an industry [is described as natural monopoly]. . . . The reason is that fixed costs are very large in relation to demand. If they can be spread over the market's entire output, a single firm supplying that output may have a lower average cost of production than two equally efficient firms, each of which would incur the same fixed costs but be able to spread them over only one-half the output.[272]

> [3] A natural oligopoly exists where the entire demand for a good or service can be satisfied at lowest cost by only a few firms . . . . A natural monopoly exists when the entire demand for a good or service can be satisfied at lowest cost by one firm.[273]

---

271. Yoo, *supra* note 4, at 849.

272. Sandeep Vaheesan, *Reviving an Epithet: A New Way Forward for the Essential Facilities Doctrine*, 2010 UTAH L. REV. 911, 911 (2010) (alterations in original) (quoting RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 367 (7th ed. 2007)).

273. RICKS ET AL., *supra* note 15, at 8–9.

There are four ways in which these definitions retain too much of the outdated neoclassical model of competition and take too little from the imperfect competition revolution.

First, the definitions do little to acknowledge the difficulty of drawing a line between the "ordinary" economies of scale enjoyed by many workably competitive firms in our economy and the "extraordinary" economies of scale that are supposed to justify natural monopoly. Firms don't wear their "U" or "L" shaped cost curves on their foreheads, and never have. In Alfred Marshall's time, however, it was at least somewhat plausible that a few sectors with exceptional economies of scale would stand out from the large pack of perfectly competitive industries, allowing us to identify them as natural monopolies. In our imperfectly competitive world, by contrast, many firms benefit from significant economies of scale, so the distinction between a natural monopoly and a workably competitive oligopoly is exceedingly subtle and fine-grained in a way the definitions fail to acknowledge.

Second, the definitions propose to identify natural monopoly based on the cost of production alone, but this ignores the insights of Chamberlin, Robinson, and Schumpeter about the possibility of product competition.[274] If we look only at production cost curves alone, we are likely to conclude that many goods and services are most efficiently produced by a single integrated enterprise. Indeed, almost all software meets the definitions of natural monopoly stated above: most of the cost of developing and improving software is fixed, and the marginal cost of making it available to each additional user is often nearly zero, generating powerful economies of scale. But there is more to the story. Product differentiation and dynamic innovation may well allow competition to exist even in the face of inexhaustible economies of scale. One video game, for example, will not conquer the market, in spite of near-infinite economies of scale—we want variety.

Third, the definitions suggest that our goal is to minimize the cost of production, but the imperfect competition revolution taught us that we must inevitably make tradeoffs between the goal of productivity and the goal of efficient allocation of resources through competition. Even when monopoly would minimize production cost, we might nevertheless be better off with oligopolistic competition.

Finally, our concept of natural monopoly does not have enough to say about "new economy" phenomena. When we look at the new technologies of our time, we often conclude that the market power we are most concerned with is not related to costs of production but to network effects (which make a product or service more valuable when more people use it) or first-mover

274. *Supra* Part I.E.

advantages (such as the lock-in of standards).[275] Further, we understand that the monopolizing force of those phenomena is sometimes defeated by countervailing phenomena like interoperability, multi-homing, and the rapid pace of innovation—all of which create spaces for competition.[276] The definitions of natural monopoly do not tell us what to do with these insights. They remit us to the question of production cost alone.

B. An Updated Concept of Natural Monopoly

To address these shortcomings, I propose the following definition for natural monopoly. My definition is intentionally pragmatic, not ontological, in that it collapses the question of whether an asset *is* a natural monopoly into the question of what we should *do* (regulate or not). An asset or service should be regulated as a natural monopoly only if all five of the following criteria are met:[277]

(1) Inexhaustible economies of scale, due to either:

  (a) declining average cost or

  (b) network effects without interoperability or multi-homing;

(2) Low product differentiation;

(3) Vertical severability from non-natural monopoly assets;

(4) Sufficient tenure as an apparent natural monopoly; and

(5) Regulation serves a compelling social purpose.

---

275. *See, e.g.*, *Antitrust and Platform Monopoly, supra* note 8, at 1962–63.

276. *Antitrust and Platform Monopoly*, *supra* note 8, at 1974–76 (explaining the role of interoperability and multihoming); Adam Thierer, *The Perils of Classifying Social Media Platforms*, 21 COMMLAW CONSPECTUS 249, 250–51 (2012) (arguing against regulation of social media platforms on the grounds that innovation will ensure sufficient competition).

277. In *Antitrust and Platform Monopoly*, Herbert Hovenkamp provides "five related factors [that] determine the existence of a natural monopoly," on which my test is substantially based. *Antitrust and Platform Monopoly*, *supra* note 8, at 1972. Elements 1 and 2 of my test re-state his points 3–5 but reorganize them to recognize that interoperability and multihoming are an "antidote" to network effects and scale. Hovenkamp's discussion of his points 1 and 2 ("lack of stable competition" and "durability of a dominant position and the ability to accommodate or resist technological change") suggest they are better understood as effects or indicia of natural monopoly, not a necessary or sufficient element thereof. I replace them with criteria that I think better capture the inquiries needed to designate an *asset* (as distinct from a firm) as a natural monopoly.

This definition, like its neoclassical predecessors, conceptualizes natural monopoly as a peculiar relationship between a technology of supply and the optimal structure of the market for the goods or services produced by the technology. However, it is "thicker" than the neoclassical approach, supplementing the neoclassical definition's abstract and conclusory notion of "subadditivity" with several institutionalist and historical elements.[278] In the remainder of this section, I explain each of the test's five elements and apply them to the "Big Tech" landscape.

1. Inexhaustible Productive Economies of Scale or Network Effects

Inexhaustible economy of scale remains the foundation of natural monopoly, as it was in the neoclassical system. The updated definition, however, recognizes that economies of scale can come from either of two sources. They may come in the traditional way: a high ratio of fixed costs relative to marginal costs can result in an average cost curve that slopes down throughout the relevant range of potential outputs, without ever turning upwards into a "U" shape. Alternatively, network effects can result in a product whose value to customers increases with the percentage of the market the product serves. For example, a telephone system that serves everyone is more valuable than a telephone system that serves only one of every three people, with the other two using different systems. Network effects are sometimes described as something different from economies of scale, but from a lexical perspective, they are just as well classified as a subtype thereof. Productive economies are a supply-side economy of scale; network effects are a demand-side economy of scale. In both cases, bigger is more economical. In this way, network effects play the same functional role in the natural monopoly inquiry as productive economies: they initiate the analysis by giving us a reason to believe that there is a "Blessing of Bigness" that we need to explore.

Network effects, however, can be defeated by interoperability or multi-homing. For example, in spite of the network effects that favor the existence of just one telephone network, telephone service providers like AT&T, Sprint, Verizon, and T-Mobile are not natural monopolies, because they can interoperate with one another. It is as easy for a Verizon customer to call an

278. There is, perhaps, some poetic justice in this. The earliest concepts of natural monopoly were created by lawyers and economists inspired by institutionalist and historicist critiques of British Political Economy's abstractions. Richard Ely, for example, listed five factors (different from mine) in his 1889 summary of Telegraph Monopoly. *Telegraph Monopoly*, *supra* note 117, at 49.

AT&T or Sprint customer as it is to call another Verizon customer.[279] Similarly, a single ride-hailing market that connects every potential driver to every potential rider would be more valuable than a ride-hailing platform that connects only half the drivers to half the riders, as might be the case if the market is split between two providers like Uber and Lyft. But as long as each driver and each rider can multi-home by using both the Uber and Lyft apps, and efficiently switch between them on their phones, ride-hailing will operate substantially as one network and will not qualify as a potential natural monopoly under the first prong of our test.[280]

In sum, the first prong of the natural monopoly test preserves the essence of the original concept (economies of scale) but integrates into it the technological phenomena of greatest relevance to the industrial organization problems of our era, thereby connecting these phenomena to their approximate historical analogues.

2. Low Product Differentiation

The low product differentiation criterion embodies one of the key insights of imperfect competition: the possibility that product competition can preserve a place in the market for multiple competitors even in the presence of significant economies of scale. Consider software. Judged by the economy-of-scale criterion alone, virtually every software product would need to be classified as a natural monopoly. The investment to write the code is fixed, and the marginal cost of making it available to an additional user is near zero. Yet no one would argue that every software product is a natural monopoly, because many are highly differentiated from each other. Think of video games or movies: high product differentiation leads to ferocious competition even when economies of scale are very large.

Indeed, it is easy to imagine many internet technology markets with partial product differentiation settling into workable oligopolistic competition of the type that prevailed in the industrial economy of the 20th century: Uber and Lyft, DoorDash and UberEats, Spotify and Apple Music, Netflix and Amazon Prime Video, Facebook and TikTok, Amazon Web Services and Microsoft Azure, Amazon.com and Walmart.com. These oligopolists may engage in anticompetitive conduct—as the Department of Justice (DOJ) has alleged—but such conduct can be addressed by antitrust law without the need for natural monopoly regulation. Thus, the second

---

279. Herbert Hovenkamp, *Antitrust Interoperability Remedies*, 123 COLUM. L. REV. 1, 5 (2023) [hereinafter *Interoperability Remedies*].

280. *See Antitrust and Platform Monopoly*, *supra* note 8, at 2035 (explaining the concept of multihoming).

criterion of our test reminds us that when we look for a natural monopoly, we are looking for exceptions to workable competition achieved by product differentiation—a rare combination of highly scalable unit economics and low differentiation.

3. Vertical Severability from Non-Natural Monopoly Functions

As described above, one of the most important learnings of the deregulatory era is that the boundaries of a natural monopoly do not necessarily coincide with the boundaries of the business built around it. The innovative firms that first brought electricity service and telephone service to market were highly vertically integrated, but we later came to believe that only some of the functions they provided were natural monopolies. The goal of any natural monopoly test must be to correctly label the *assets or functions* that qualify as natural monopolies, not to label the *firms* that control those assets or functions.

Thus, to apply our natural monopoly test to the GAFA companies (Google, Apple, Facebook, and Amazon) that have been most subject to suspicion in Big Tech literature, we must judge not the enterprises themselves nor their product lines, but rather particular assets and functions they use to create these product services. For example, judged as a service, internet search is probably not a natural monopoly, because it is subject to differentiation: different providers can compete to offer algorithms that return differentiated search results, or better visual presentation of results, or less advertising. By contrast, the web index on which the search must run is a relatively undifferentiated database with very high economies of scale. Google indexes about 45 billion webpages.[281] Microsoft's Bing search engine, the only other English-language search engine that maintains its own web index, indexes only 7 billion webpages—a feat that is said to have required an investment of $4.5 billion.[282] Moreover, crawling the web to read webpages is costly and disruptive to websites in a way that favors a single crawler. The House Subcommittee Report explains:

> Today several major webpage owners block all but a select few crawlers, in part because being constantly crawled by a large number of bots can hike costs for owners and lead their webpages to crash. The one crawler that nearly all webpages

281. *The Size of the World Wide Web (The Internet)*, WORLDWIDEWEBSIZE.COM, https://www.worldwidewebsize.com/ (last updated Jan. 15, 2025).

282. SUBCOMM. ON ANTITRUST, COM., & ADMIN. L., HOUSE COMM. ON JUDICIARY, 117TH CONG., INVESTIGATION OF COMPETITION IN DIGITAL MARKETS 147 n.1061 (Comm. Print 2020).

> will allow is Google's 'Googlebot,' as disappearing from Google's index would lead most webpages to suffer dramatic drops in traffic and revenue. Any new search engine crawler, by contrast, would likely be blocked by major webpage owners unless that search engine was driving significant traffic to webpages—which a search engine cannot do until it has crawled enough webpages.[283]

This situation is analogous to the technological conditions that make it more efficient for homes to be connected to just one local "poles and wires" infrastructure for electricity than to multiple duplicative physical infrastructures.

Facebook is similar. The front-end features of social networking have a high degree of potential product differentiation: how should content be ranked? How presented? How published? How filtered? However, the database of self-published content, which is an input to the overall social networking service, may turn out to satisfy the first two prongs of the natural monopoly test: it is relatively undifferentiated and has significant economies of scale driven by the network effects of being able to connect the most producers of content with the most users.

If these features can indeed be shown to meet the other prongs of the natural monopoly test, then there *may* be a case for separating the assets from non-natural monopoly functions and making their services accessible to all competitors on reasonable and nondiscriminatory terms.[284] That is, user posts to services like Facebook, Instagram, TikTok, and the like can be communally shared, which maximizes network effects while encouraging competition on the front-end user experience. Obviously, however, the application of such an invasive remedy must be approached with great humility. There may well be sound technological and business justifications for the vertical integration of a natural monopoly asset with related non-natural monopoly assets. This, too, is an essential lesson of the deregulatory era: forced interoperability of the telephone system led to dramatic improvements in service and cost efficiency, but forced interoperability of

283. *Id.* at 63.

284. We are speaking here of a "reverse Baxter doctrine." Joskow & Noll, *supra* note 224. Facing a regulated natural monopoly, William Baxter disintegrated its true natural monopoly functions (local telephony) from the potentially competitive segments (long distance, telephone equipment) and *removed* public utility regulation from the potentially competitive segments. Facing unregulated Big Tech enterprises, we can disintegrate their natural monopoly functions from their potentially competitive functions, and *impose* public utility regulation on the spun-off natural monopolies, to offer services to a competitive market at nondiscriminatory rates.

the electricity grid led to periodic coordination crises without any noticeable improvement in service or cost efficiency.

The point of the third prong of our natural monopoly test is to insist that before we vertically dismember a business on the theory that part of it is a natural monopoly, we must take account of the potentially valid justifications for integration. For example, if we think Google's web indexer and Facebook's self-published content database may be a natural monopoly, we must ask whether they can be efficiently separated from the "front end" search engine and social media platform, respectively. This is a fact-intensive question that demands technical analysis beyond the scope of this Article, though it has been undertaken to some extent by recent work on interoperability policy.[285]

4. Sufficient Tenure as an Apparent Natural Monopoly

The fourth criterion requires sufficient tenure as an apparent natural monopoly and thereby implements the insights of Schumpeter, Romer, Friedman, and Posner—namely, that the *least bad* alternative may be to tolerate some monopoly for some amount of time, hoping that innovation will erode it, and stepping in with government intervention only when we have watched the industry for some time and are confident in our diagnosis.[286] This criterion, therefore, imposes some additional humility on our analysis of natural monopoly.

Indeed, our worst fears often go unrealized. Microsoft, whose dominance of the computer operating system market seemed so threatening at the time of the DOJ antitrust action in the late 1990s, saw its market share eroded by competition in the decade thereafter.[287] Articles written just a few years ago include companies like Uber and Airbnb on lists of threatening "internet giants," but time has made these businesses seem more conventional.[288] Five years from now, the concerns I express in this article about Google's and Facebook's power may seem similarly passé. Indeed,

---

285. *See, e.g.*, *Equitable Interoperability*, *supra* note 29, at 1018 ("We have engaged in conversations with industry participants and technical experts about the difficulty and cost of carrying out interoperability from a technical perspective.").

286. It also follows the approach of some of the earliest theorists of natural monopoly. Richard T. Ely, Henry C. Adams, and Charles Francis Adams, Jr. were all inspired by the German Historicist school of economics and all supported their theories of natural monopoly with decades of observations regarding the history of the railroad and telegraph industries. *First Great Law & Economics Movement*, *supra* note 110, at 997, 1021.

287. Richard Blumenthal & Tim Wu, Opinion, *What the Microsoft Antitrust Case Taught Us*, N.Y. TIMES (May 18, 2018), https://www.nytimes.com/2018/05/18/opinion/microsoft-antitrust-case.html.

288. *See The New Utilities*, *supra* note 16, at 1669.

that may have already happened during the year that this article has been in press: the emergence of artificial intelligence products has fundamentally altered the market for internet search.[289] Even in scenarios where problematic market power persists over a significant period of time, patience is an aid to diagnosis of its cause: a few years of delay will often do much to reveal whether a monopoly was gained through anticompetitive conduct (in which case antitrust is the solution) or through low product differentiation and economies of scale (in which case natural monopoly regulation may be in order).

Moreover, as Schumpeter argued, temporary monopoly rents may be a feature of our economic system, not a bug.[290] Such rents may be part of the incentive that encourages innovators to bring useful new products and services into the world. If this seems inherently noxious, consider that our patent system follows a similar theory: we award property rights in ideas as an incentive to innovation. Tolerating monopoly as an incentive to innovate is little different, as long as the duration of the monopoly is, like a patent, temporary.

5. Social Purpose

Even when a firm meets the above four factors, the imposition of public utility regulation remains dangerous. We will seldom be sure of our judgment. Perhaps there is enough potential product differentiation to facilitate some degree of competition. Or perhaps the dynamics of innovation will render the function irrelevant within a few years, as cellular technology did to local telephone poles and wires. Moreover, given what we know about the cost of regulation—both direct and in the form of diminished competition—we will be even less confident that the benefits of regulation outweigh its costs, especially since we are talking about reaching inside businesses to regulate particular assets. For that reason, we should stay the hand of regulation unless it is demonstrably justified to advance important social objectives.

This fifth prong, therefore, ensures that invasive public utility regulation is not applied to "ordinarily" oligopolistic or even to all naturally monopolistic markets. If it turns out that the soft drink production sector meets all of the four factors above—in other words, it has massive economies of scale, limited product differentiation, can be cleanly severed from all

289. *See* United States v. Google, No. 1:20-cv-03010-APM, slip op. at 1 (D.D.C. Sept. 2, 2025) (describing how "the emergence of GenAI changed the course of this case" in the year between the end of the liability opinion and the remedies order)

290. *Supra* Part II. E.

related industrial sectors, and has been controlled by an oligopoly or monopoly for years—but the only proffered objective of regulation would be to marginally reduce beverage prices, it would still not qualify for natural monopoly regulation. Indeed, in my view, economic efficiency alone will almost never be enough to justify such regulation.

By contrast, the cultivation of competition in social networking and internet search has a well-demonstrated social purpose; indeed, it is a cause with world-historical implications. We have all come to understand the importance of Facebook's and Google's decisions about how to rank content and about which users and posts to exclude from each platform.[291] Indeed, as Robert Epstein has shown, Google could tip the results of an election simply by changing the ranking of results in its search.[292] Jonathan Zittrain has pointed out that merely changing the "doodle" on Google's search landing page could have similar effects.[293]

Direct government action to curtail such power is worse than the disease. On foundational First Amendment principles, it is even more problematic for the government to exercise editorial power over the news than for a private monopoly to do so. The government's jawboning of social media platforms during the pandemic arguably led to the suppression of legitimate points of view about the origins of the COVID-19 virus, vaccine efficacy, and the contents of Hunter Biden's laptop.[294] Anti-"censorship" statutes passed by Texas and Florida, if allowed by the courts, would only take government further into this field. These statutes would deprive social media of the type of editorial control that newspapers routinely exercise, and inevitably require the government or courts to draw lines between forbidden "censorship" and allowable exclusion of inappropriate conduct, or at the very least, judge the consistency of the platforms' application of their own editorial standards.[295] In 2025, the federal government enacted a law forcing the Chinese owners of the TikTok platform to divest their U.S. operations.[296] This may reduce the

---

291. *See* Evelyn Douek, *Content Moderation as Systems Thinking*, 136 HARV. L. REV. 526, 531 (2022) (describing content moderation as a "vast system of administration that includes a far broader range of decisions and decisionmakers than the standard picture admits").

292. *Why Google Poses a Serious Threat to Democracy, Testimony Before the U.S. S. Judiciary Subcomm. on the Const.*, 116th Cong. 2 (2019) (testimony of Robert Epstein, Ph.D., Senior Research Psychologist, American Institute for Behavioral Research and Technology).

293. Jonathan Zittrain, *Engineering an Election*, 127 HARV. L. REV. F. 335, 337 (2014).

294. Murthy v. Missouri, 144 S. Ct. 7, 8 (2023) (Alito, J., dissenting).

295. *See* Moody v. NetChoice, LLC, 144 S. Ct. 2383, 2394 (vacating and remanding to the Eighth and Eleventh Circuits for proper analysis of the facial First Amendment challenges to the Florida and Texas laws).

296. David McCabe, *TikTok Flickers Back to Life After Trump Says He Will Stall a Ban*, N.Y. TIMES (Jan. 19, 2025), https://www.nytimes.com/2025/01/19/technology/trump-tiktok-ban-executive-order.html.

threat of Chinese interference in U.S. elections, but it raises the specter of other forms of interference: the U.S. government might steer the sale to a favored buyer, in exchange for favorable coverage or other political advantage.

Competition of the type that could be unlocked by the opening of back-end natural monopolies to competing front-ends is a better solution to the problems posed by social media companies' market power. If there were five social media platforms, each with the ability to access all self-published content, their content moderation decisions would be little more problematic than the New York Times' publication of a story that the Wall Street Journal finds unnewsworthy. Similarly, if there were three major search engines, bias by one of them in search results would be relatively easy to detect and avoid by switching to an unbiased competitor.

Therefore, if it can be shown that Google's web indexer or Facebook's database of self-published content meets the other four criteria for natural monopoly regulation, there is a strong case that their regulation as public utilities would be justified by a sufficiently compelling social purpose.

C. Application to Contemporary Problems of Industrial Organization

In the course of illustrating the five factors of my proposed test, I have already substantially described what I think is its most immediate and important application: the problem of Big Tech market power.[297] Put simply, my contention is that the use of existing antitrust doctrine to improve competition in oligopolistic tech markets probably addresses most of the Big Tech Problem—but not all of it. The part of the problem that remains, I think, is more or less coextensive with the old natural monopoly problem: private enterprises control difficult-to-reproduce gateway functions that are important inputs to downstream economic activity. Now, instead of inventing new words or frameworks to describe the problem, contorting antitrust doctrine to include new theories of liability, or imposing ill-targeted interoperability mandates—the problems identified in Part II above—we should update our understanding of natural monopoly and use it to scrutinize the Big Tech landscape for qualifying assets. We should then spin those assets off from the competitive portions of the businesses that control them

297. My concept of natural monopoly, however, is not targeted at the Big Tech Problem alone. It also has immediate applications, I think, to questions of how to reform "grid governance" in the electricity sector (described above). Additionally, as an industry-agnostic general statement of what we have learned about natural monopoly, it can be applied as a heuristic to test whether any new technology may entail a natural monopoly.

and force them to offer service to all competitors at a reasonable price on nondiscriminatory terms.

The analysis of which, if any, Big Tech assets actually qualify for this treatment is a fact-specific, technical exercise beyond the scope of this Article. However, we know enough of the facts already to exonerate many features of the Big Tech landscape that have been unnecessarily subjected to suspicion. Online marketplaces, social networking "front ends," app stores, and operating systems are differentiated products, amongst which competition is probably possible and valuable. Other products and services—such as Uber, Lyft, and DoorDash—are also probably too young or too inconsequential from a policy perspective to merit aggressive intervention.

The updated theory of natural monopoly directs our attention instead to further analyses of back-end, undifferentiated "utilities" controlled by Big Tech enterprises, such as Google's web indexer and Facebook's database of self-published content. The concept of natural monopoly has the potential to better guide us in determining which of these utilities should be regulated and what the goals of such regulation should be—a prospect that I more deeply explore in the next Part.

Such a reform would *not* be "pro-monopoly." It is better described as "anti-oligopoly." Specifically, its goal is to identify natural monopoly assets currently held by oligopolistic competitors, and open access to them at just, reasonable, and nondiscriminatory prices, thereby enabling competition in the downstream product markets that rely on the assets—including information gateways important to the health of our democracy.

D. Natural Monopoly as the Criterion for Interoperability Mandates

One of the most promising lines of thinking on the contemporary problems of Big Tech suggests that we use interoperability mandates, implemented by regulation or antitrust remedies, to force dominant firms to make certain assets or services available to their rivals for incorporation into competitive offerings.[298] The various flavors of this reform include "network neutrality," "open access," "anti-discrimination rules," prohibitions on self-preferencing, mandatory licensing at Fair, Reasonable, and Non-Discriminatory (FRAND) rates, "equitable interoperability," and "joint

298. *See, e.g.*, *Equitable Interoperability*, *supra* note 29, at 1016; *Interoperability Remedies*, *supra* note 279, at 1; Tim Wu, *Why Have a Telecommunications Law? Anti-Discrimination Norms in Communications*, 5 J. ON TELECOMM. & HIGH TECH. L. 15, 28 (2006) [hereinafter *Anti-Discrimination Norms*].

management of unified productive assets."[299] The District Court for the District of Columbia recently gave these ideas additional importance by ordering an interoperability mandate in the remedies phase of the Google Search antitrust proceeding. Specifically, the Court ordered Google to make certain web indexing data available to competitive search providers.[300]

These ideas, I think, are on the right track. The literature on interoperability, unlike that on the revitalization of antitrust, is alert to the relationship between the internet's special characteristics as a technology of supply and the problematically concentrated outcomes in the industrial sectors currently controlled by Big Tech enterprises.[301] It correctly focuses attention on particular assets or functions, rather than on a broad range of unsensational conduct. And it recognizes that forcing dominant firms to give their competitors access to these assets or functions on a nondiscriminatory basis can unlock competition of the kind we most want. For example, it envisions competitive search engines built on a shared back-end "web indexer" service, each vying to create the best search algorithms and front-end presentation. Similarly, it envisions competitive social networks that share self-published content via back-end APIs whilst competing to offer their users the best "front end" by which to access that content.[302] The potential salutary effects of such competition include diversification of approaches to content moderation and ranking of results, diversification of business models (ad-supported, subscription, and hybrids thereof), and reduction in the "attentional price tag" charged by the ad-supported services. This type of competition is more compelling than what would likely arise from the dismemberment of these enterprises into smaller but functionally identical "Micro Googles" or "Baby Books," or the separation of business units like Amazon's marketplace from its web store—interventions that reduce efficiency and dampen valuable network effects without an obvious corresponding advantage.[303]

---

299. Tim Wu, *Network Neutrality, Broadband Discrimination,* 2 J. ON TELECOMM. & HIGH TECH. L. 141, 141–42 (2003) (describing network neutrality and open access); *Anti-Discrimination Norms*, *supra* note 298, at 28 (anti-discrimination rules); American Innovation and Choice Online Act, S. 2992, 117th Cong. (as reported by S. Comm. on the Judiciary, Mar. 2, 2022) (prohibiting self-preferencing that materially harms competition); *Equitable Interoperability*, *supra* note 29, at 1016 (equitable interoperability); Paul Heidhues et al., *More Competitive Search Through Regulation*, 40 YALE J. ON REGUL. 915, 945 (2023) (mandatory licensing at FRAND rates); Hovenkamp, *Antitrust and Platform Monopoly*, *supra* note 8, at 2022 (joint management of unified productive assets).

300. United States v. Google, No. 1:20-cv-03010-APM, slip op. at 146–47 (D.D.C. Sept. 2, 2025).

301. *Separation of Platforms and Commerce*, *supra* note 7, at 1076–77; *Equitable Interoperability*, *supra* note 29, at 1015.

302. *See Equitable Interoperability*, *supra* note 29, at 1016.

303. *Interoperability Remedies*, *supra* note 279, at 1.

The criteria that should trigger interoperability mandates are, however, under-theorized. A typical justification for intervention is the network effects said to be endemic to "large digital platforms."[304] But no one thinks that *all* assets held by such platforms should be shared. An important recent article, for example, proposes the sharing of self-published content amongst social networks like Facebook, plus prohibitions on discrimination and self-favoritism by Apple's and Google's operating systems and app stores, as well as by Amazon's marketplace.[305] It is difficult to discern the specific criteria that unite this group of regulatory targets, and differentiate it from other technology assets that escape scrutiny. Why Google's web indexer but not its Gmail server or search algorithms? Why Google's operating system but not Microsoft's? For that matter, are we sure that "network effects" alone justify treating Big Tech differently than other industries? If Facebook invests $1 billion in a social networking platform and Tesla invests the same amount in a new electric vehicle chassis design, why should Facebook be compelled to open its platform to competitors while Tesla is not? Both scenarios implicate the same trade-off between incentivizing innovation and facilitating competition, so the fact that only one investment (Facebook's) is said to have network effects is not a satisfying rationale for their differential treatment. In short, the principle that qualifies specific assets and services for interoperability mandates and disqualifies others remains fuzzy.

These same questions haunt the extensive commentary on the "antitrust version" of interoperability: the controversial theory that a "unilateral refusal to deal" or denial of access to an "essential facility" might constitute anticompetitive conduct.[306] This theory of liability was recognized (equivocally) by the Supreme Court in *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*,[307] condemned by antitrust scholars for its tendency to encourage free riding on competitors' investments,[308] and then narrowed by

304. *Id.* at 33. In keeping with this view, the American Innovation and Choice Online Act would apply non-discrimination rules to all "online platforms" of a certain size. *See* American Innovation and Choice Online Act, S. 2992, 117th Cong. § 2(a)(9), 3(a) (as reported by S. Comm. on the Judiciary, Mar. 2, 2022).

305. *Equitable Interoperability*, *supra* note 29.

306. Robert Pitofsky et al., *The Essential Facilities Doctrine under U.S. Law*, 70 ANTITRUST L.J. 443, 445 (2002) (describing the history of the doctrine and debates about its proper scope).

307. *See* 472 U.S. 585, 600–11 (1985) (recognizing in a narrow set of circumstances that a monopolist's unilateral refusal to continue a prior course of dealing with a rival, undertaken without efficiency justification and at the expense of short-term profits, can support Section 2 liability). The doctrine also has roots in United States v. Terminal R.R. Ass'n, 224 U.S. 383, 411 (1912).

308. *See, e.g.*, Phillip Areeda, *Essential Facilities: An Epithet in Need of Limiting Principles*, 58 ANTITRUST L.J. 841, 849–50 (1990) (expressing concerns about the potential over broadness of the essential facilities doctrine).

Verizon Communications, Inc. v. Law Offices of Curtis V. Trinko[309] to the circumstance in which a monopolist inexplicably breaks off a previous course of dealing. Recently, it has been the subject of dozens of academic papers that see in the doctrine a potential solution to the Big Tech Problem.[310] This commentary, however, has not hit on anything approaching a consensus regarding which unilateral refusals to deal constitute anticompetitive conduct, or which facilities are so "essential" that they must be licensed to competitors on demand. In fact, the doctrine has been described as "a good way to elicit eyerolling within antitrust circles."[311] Without a sound limiting principle, we find ourselves stuck between our aversion to forcing firms generally to give their competitors access to the proprietary fruits of their investments and the sense that there is nevertheless *something* in the essential facilities doctrine that gets closer to the heart of the Big Tech problem than anything else antitrust has to offer.

The biggest difficulty in determining which assets should be subject to interoperability mandates is often said to be the need to make trade-offs between competition *on* a platform and competition *between* platforms.[312] For example, do we want to force computer operating systems to offer equal access to all apps in order to maximize app competition on the platform? Or do we want to allow each operating system to curate and influence the apps with which it partners in order to encourage more diversity and competition between operating systems? Do we want to force a bridge over the Mississippi to interconnect with all railroad lines in order to maximize competition on the bridge? Or do we want to instead encourage each railroad to build its own bridge, in hopes that inter-bridge competition will lead to better bridges?

Natural monopoly is the missing principle that justifies the application of interoperability to the relatively easy subset of these cases. When we judge

---

309. 540 U.S. 398, 408–09 (2004).

310. Thomas Nachbar, *Essential Facilities and the Law of the Hammer*, ANTITRUST CHRON., Apr. 2023, at 1, 14 ("[T]here are many, many law review articles offering to 'revitalize,' 'revive,' 'renew[]' or otherwise resurrect the essential facilities doctrine to solve any range of competitive ills, from self-dealing by internet shopping platforms to social media platform refusals to allow application interfaces."); *The Antitrust Duty to Deal in the Age of Big Tech*, *supra* note 18, at 1385–86 ("[T]he rise of dominant platforms like Google, Facebook, and Amazon has provoked intense debate over whether the antitrust duty to deal needs a revival.").

311. Erik Hovenkamp, *Trinko Meets Microsoft: Leverage and Foreclosure in Platform Refusals to Deal*, ANTITRUST CHRON., Spring 2023, at 1, 22, 27.

312. *See, e.g.*, *Equitable Interoperability*, *supra* note 29, at 1023 (noting that interoperability can shift competition from being *for* the market to being *in* the market); *Interoperability Remedies*, *supra* note 279, at 36 ("Interoperability is a two-sided coin. One of the great values of competition, and of digital competition in particular, is its diversity. Excessive interoperability covering too many of the features of individual firms may simply serve to homogenize the market, destroying competitive incentives and inviting free riding.").

an asset to be a natural monopoly, we are saying that its duplication by competitors has virtually no value. We don't want the competing railroad companies in *United States v. Terminal Railroad Ass'n*[313] to each build their own bridge over the Mississippi or (perhaps more controversially) competing search engines to each build their own web indexer. We don't want each telephone company to invent its own network protocols; we want them to share the same protocols to function as one telephone network. In short, in the natural monopoly scenario, we are completely committed to competition *on* the platform because we perceive little value in competition *between* alternative versions of the platform.

The harder cases, which have generated so much scholarly fear, arise from the forced sharing of non-natural monopoly assets. Mandating that operating systems provide equal access to all the apps that want to compete on their platform while also encouraging some competition between operating systems is an example of a harder case. This intervention, unlike the natural monopoly cases above, entails significant potential costs as well as potential benefits: our efforts to increase app competition may end up decreasing operating system competition by constraining differentiation. The same is true for forced access to competitive marketplaces.

My proposal is that we avoid these difficult questions for now by first applying interoperability mandates only to natural monopoly assets—a category that, for the reasons described above, likely does not include operating systems, search engines, or social media platforms, but only certain special back-end assets under their control. My intuition is that solving the natural monopoly problems may be enough: we may find that our efforts have already encouraged enough downstream competition amongst the Big Tech products and services built on top of the newly opened natural monopolies to leave the non-natural monopoly assets well enough alone.

The idea that natural monopoly might serve as the limiting principle for essential facilities sharing and interoperability mandates is not new.[314] The reason it has not been more widely accepted, I think, is the sorry state of our natural monopoly concept. The scholars who see potential in the essential facilities doctrine are more interested in network effects and related "new economy" phenomena than in the "L"-shaped production cost curves described by the neoclassical definition of natural monopoly.[315] Given that

313. 224 U.S. 383 (1912).

314. Herbert Hovenkamp, *Federal Antitrust Policy: The Law of Competition* § 7.7a, at 402 (6th ed. 2020) ("most of the things found by courts to be essential facilities have [been] . . . natural monopolies" or assets regulated or provided by the government).

315. *See, e.g.*, Nicholas Guggenberger, *Essential Platforms*, 24 STAN. TECH L. REV. 237, 276–85 (2021) (providing a number of such justifications for essential facilities sharing without mentioning

they see natural monopoly as a confused and inapposite phenomenon, it stands to reason that they think the concept does not define circumstances in which interoperability should be applied. My hope is that the updated test for natural monopoly I presented above can better explain the true nature and scope of the natural monopoly problem and thereby reveal its potential to serve as the guiding principle for interoperability mandates.

E. But Do We Really Need Public Utility Regulation?

In this final Part, I address an important objection. Assuming we agree that natural monopoly is the problem, do we need to apply public utility regulation as traditionally understood, with all the known downsides it entails? Is there a less invasive means of achieving our goals, such as by the application of a minimalistic nondiscrimination rule or lightweight interoperability mandate?

These questions reflect deep suspicions about public utility regulation that first emerged in the 1960s on both the left[316] and right,[317] and have now assumed the status of conventional wisdom. The "left" version presented itself as a revisionist history of the Progressive Era, critiquing the extent to which institutions, like public utility regulation, were "captured" by big business.[318] The "right" version presented as a generalized positive theory of government action, often referred to as "public choice" theory.[319] The idea is that the facts of economic regulation do not fit the theory that publicly interested government actors impose regulation in order to correct market failure. The benefits of such a correction are diffuse, making it difficult for dispersed potential beneficiaries to organize themselves to promote political action. Instead, the theory goes, economic regulation is sought out by companies that seek the "quiet life of a monopolist" and use a smokescreen of public interest rhetoric to disguise their agenda.[320] The regulation that results is "worse than ineffective"; it not only fails to protect the public from monopolistic prices but also dampens innovation and investment because the

---

natural monopoly); *The Antitrust Duty to Deal in the Age of Big Tech*, *supra* note 18, at 1491 n.31 (recognizing that essential facilities sharing is appropriate for "exceptional" natural monopoly scenarios but treating that as an exceptional circumstance, not the core essential facilities scenario); Vaheesan, *supra* note 272, at 912–13 (arguing that open access regimes ordered by public utility regulators adequately cover tangible natural monopoly assets, so the essential facilities doctrine should focus on intangible essential facilities, which may or may not be natural monopolies).

316. GABRIEL KOLKO, RAILROADS AND REGULATION, 1877–1916 (1970).

317. MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS (1965).

318. KOLKO, *supra* note 316, at 233–34.

319. OLSON, *supra* note 317.

320. *Id.*

natural monopolist finds ways to enroll regulators in its efforts to exclude upstart technologies.[321]

One consequence of these views has been a generalized reluctance to extend regulation to new technology. The quest for a minimalistic alternative to regulation is a hallmark of internet policy over the thirty or so years of its existence. "Net neutrality," for example, was conceived by its architects as "a light form of behavioral regulation that narrowly targets the behavior identified as problematic and is far less intrusive than other forms of regulation."[322] In the subsequent Big Tech debate, the weight of academic opinion tends to lean instead towards nondiscrimination or interoperability remedies similar to the net neutrality concept, again conceived as less invasive than traditional regulation and hopefully implemented via antitrust-style adjudication, not regulation.[323] More radical positions have emerged, but these have mostly centered on the idea of horizontal breakups or vertical line-of-business restrictions, without traditional regulation of the unbundled lines of business that would result from such reforms.[324] Indeed, even the writers who draw on the public utility tradition tend to advocate only for reclaiming the Progressive "ethos" of public utility thinking, disclaiming any support for the institutions and techniques of public utility regulation.[325]

---

321. *See, e.g.*, THE MASTER SWITCH, *supra* note 222, at 55; *Amazon's Antitrust Paradox*, *supra* note 7, at 800; *Natural Monopoly and Its Regulation*, *supra* note 1, at 622.

322. Barbara van Schewick, *Towards an Economic Framework for Network Neutrality Regulation*, 5 J. ON TELECOMM. & HIGH TECH. L. 329, 333–34 (2007); Wu, *Network Neutrality, supra* note 299, at 142; *Broadband Debate: A User's Guid*e, *supra* note 253, at 73; *Why Have a Telecommunications Law?*, *supra* note 298, at 17–18 (reimagining the whole of telecommunications policy around the principle of non-discrimination, applied as an "ex ante rule with ex post remedies").

323. *Amazon's Antitrust Paradox*, *supra* note 7, at 800–01 (describing application of the essential facilities doctrine to Amazon's marketplace as a "lighter version" alternative to public utility regulation); *The New Utilities*, *supra* note 16, at 1654–55 (arguing against "full imposition of traditional public utility regulation" on the grounds that it "might ratify high-priced monopolies, insulating them from competition"); *Equitable Interoperability*, *supra* note 29, at 1015, 1017 (praising interoperability as a "minimal regulation" and "a light-touch regulatory governance scheme"); *Antitrust and Platform Monopoly*, *supra* note 8, at 1971, 2050 (arguing that "[r]egulation . . . entrenches existing technologies and, in doing so, bolsters existing incumbents" and advocating shared ownership to achieve interoperability of key assets); Adam Thierer, *The Perils of Classifying Social Media Platforms*, 21 COMM. L. CONSPECTUS 249, 250–51 (2012) ("Treating these nascent digital services as regulated utilities would harm consumer welfare because public utility regulation has traditionally been the archenemy of innovation and competition.").

324. *Separation of Platforms from Commerce*, *supra* note 7, at 1078, 1083; *Amazon's Antitrust Paradox*, *supra* note 7, at 798–99.

325. William Boyd, *Public Utility and the Low Carbon Future*, 61 UCLA L. REV. 1614, 1708–10 (2014) (arguing for a "revitalized concept of public utility" that "cannot simply adopt the older concept of public utility" but needs "new ideas and conceptual innovations"); *The New Utilities*, *supra* note 16, at 1687–88 (noting that "this Article does not suggest, and is not meant to suggest, that we should mechanically copy and reinstate old models of public utility regulation" such as the "tired, old top-down institutional forms we might associate with early twentieth century rate regulation"); *Infrastructural*

This hope of replacing public utility regulation with a more minimalistic alternative is, however, a mirage. But for a few remarkable exceptions that can be solved through protocols alone, any interoperability system that works will be practically indistinguishable from traditional natural monopoly regulation. In fact, properly understood, interoperability mandates, essential facilities sharing, and public utility regulation are all the same thing—a requirement that the goods or services produced by a natural monopoly asset be made available to all interested parties at a just, reasonable, and non-discriminatory price. To support this contention, I will analyze three common flavors of "minimalistic interoperability," and then compare them to traditional public utility regulation.

1. Nondiscrimination Rules

The first variety of interoperability regime is a nondiscrimination or non-self-preferencing rule, such as net neutrality. This variety is the purest embodiment of the "lightweight regulation" dream, in that it requires subject firms to do no more than they have already done for someone else. If a firm has licensed its technology to one entity, nondiscrimination requires only that it license it to others on similar terms. The reach of these regimes, however, is correspondingly limited. They only touch assets that have already been at least partially opened to competitors (such as Amazon's web store or Apple's App Store). Thus, they leave untouched some of the most interesting potential targets of interoperability regimes, including those identified above (Google's web indexer or Facebook's database of self-published content). They also have the strange consequence of forcing a business that has once adopted an open business model to persist in that business model forever, while allowing businesses that are more consistent and scrupulous in "walling their garden" to evade scrutiny. Thus, if Amazon had remained a web retailer without ever opening its marketplace to third-party sellers, it would be above suspicion, but because it *did* allow third party sellers, it is potentially subject to a mandate that forces it to extend and perpetuate that business model. It is difficult to articulate a principled justification for punishing experimentation in this way, or for tolerating "fully closed" natural monopolies but not "partially open" ones.

Regulation, *supra* note 16, at 938 (describing the need for the ideas of public utility regulation to be "adapted and updated"); *Regulating Informational Infrastructure, supra* note 261, at 247 (explaining how the "familiar problems of regulatory capacity and capture" undermine public utility regulation).

2. Interoperability Mandates

More full-fledged interoperability regimes, such as "open access" or "must interconnect" policies, have the power to impose affirmative obligations on assets that have never been previously "opened." A good example is the court's recent order during the remedies phase of the Google search monopolization trial, which may force Google to offer access to its search index data at marginal cost rates.[326]

Interoperability can sometimes be achieved merely through mandatory protocols or interconnection requirements—a happy but relatively rare scenario that I think is consistently overemphasized in the literature. The canonical example is the telephone network, which consists of multiple competitors (AT&T, Verizon, Sprint, etc.) that route calls amongst their linked networks according to established protocols. In this case, the "natural monopoly asset" at issue is limited to an intangible set of protocols for transferring calls amongst different providers, together with some relatively trivial physical connection facilities. As there is no tangible or productive asset at issue, there is no issue of how to price the shared output. Accordingly, simple interconnection rules without traditional public utility regulation are enough.[327] In my preferred vocabulary, this situation is merely a special case of regulated natural monopoly, in which the just, reasonable, and non-discriminatory price of the asset in question is zero.

In most cases, however, there *is* a productive asset at issue—a bridge, an electricity dispatch system, a web indexer—that was created at non-zero cost to its owner and must continue to be operated, maintained, and updated from time to time.[328] This raises the question of the price and terms on which service from the asset will be made available. Consider the Google search example. The District Court for the District of Columbia ordered Google to make its data available at "marginal cost," and set up a technical committee to help implement this data sharing (among other remedies).[329] How is marginal cost to be measured? What data must be shared, specifically? And

---

326. United States v. Google, No. 1:20-cv-03010-APM, slip op. at 147 (D.D.C. Sept. 2, 2025). Similar proposals have been described in regulatory toolkits developed by coalitions of academics. Heidhues et al., *supra* note 299, at 945 (noting mandatory licensing at FRAND rates).

327. However, this interoperability regime was in fact constructed by a public utility regulator (the FCC) as part of the process of restructuring communications regulation, which may not be mere coincidence.

328. *See Interoperability Remedies*, *supra* note 279 (explaining that a survey of interoperability situations shows that most involve assets of the type mentioned).

329. United States v. Google, No. 1:20-cv-03010-APM, slip op. at 6, 146–70 (D.D.C. Sept. 2, 2025).

what other technical requirements may Google impose as prerequisites to interconnection with its indexer?

Anyone with experience in the major public utility regulatory controversies of the last fifty years will immediately grasp the difficulties in administering such a forced data-sharing regime.[330] As long as Google continues to own the asset, it will have strong incentives to inflate marginal cost and exaggerate the technical dangers of interconnection, in order to retain its advantage over its competitors. Thus, Google will offer plausible accounts of why the marginal cost of web indexing is quite high, and of the dangers that will result if interconnection with the indexer is not carefully controlled by Google via lengthy procedures and interconnection requirements. Its new competitors will offer similarly persuasive accounts for why access to the asset ought to be quite cheap and easy. Resolving these controversies will not be a "lightweight" affair. We will need a regulator—whether we assign that role to a court, agency, regulatory commission, or industry council. And we will be sending that regulator into battle with severe information asymmetry. The annals of public utility history contain similar "regulatory suicide missions," among them the above-described price-setting controversies that attended attempts to unbundle local telephone service and electricity transmission.[331]

3. Shared Governance

A third, more novel approach to interoperability is the pooling or shared ownership regimes that have recently attracted significant attention from scholars.[332] The preeminent example of this remedy is in *United States v. Terminal Railroad Ass'n*, in which a group of 14 railroads jointly owned a corporation that owned a key bridge over the Mississippi River.[333] The corporation's exclusion of other railroads from the bridge was challenged as an antitrust violation.[334] The Supreme Court ordered the reorganization of the corporation to allow the admission of other competing railroads as joint owners, essentially recognizing that the asset in question was a natural

330. *See, e.g.*, Yoo, *supra* note 4, at 851–52; Owen, *supra* note 257, at 17.

331. Owen, *supra* note 257, at 17; McNish, *supra* note 232, at 628; HOW NOT TO DEREGULATE, *supra* note 232, at 6; *but see Why Have a Telecommunications Law?*, *supra* note 298, at 43–44 (disputing "[t]he long-standing assumption . . . that any such interconnection remedy will require a complex, government administered rate-setting scheme" on the grounds that the government can "rely on the setting of prices at zero").

332. *See, e.g.*, *Antitrust and Platform Monopoly*, *supra* note 8, at 1952; *Interoperability Remedies*, *supra* note 279, at 4.

333. 224 U.S. 383, 391–94 (1912).

334. *Id.* at 394–95.

monopoly, and using shared corporate governance to make it interoperable.[335] Similar industrial structures exist in the present day and are well-known curiosities for antitrust law, including the Chicago Board of Trade (independent trading firms that compete with each other but jointly share the trading platform) and the Associated Press (a cooperative whose international news bureaus provides shared content to local newspapers).[336] Analogously, Google's web indexer or Facebook's content database might be placed under the joint control of multiple competitive search engines or social media companies. For some contemporary commentators, this approach holds out the possibility of achieving by means of private governance the same result that other interoperability regimes would obtain by government mandate, thereby avoiding the need for regulation.[337]

On closer examination, however, this approach is not so different from the other approaches, and its advantages relative to public utility regulation are not obvious. If ownership of the asset is made available to all potential competitors, including new entrants, there must be a means of determining the price and terms on which shares in the asset are sold. Accordingly, the tricky debates described above may not be avoided, but merely re-cast from questions about the terms of contractual access to an asset into questions about the terms of sale of an intangible property right (in other words, a stake in the entity that owns the asset).

Additionally, unregulated ownership of a natural monopoly asset by a small group of oligarchic competitors is not necessarily a desirable outcome. As Herbert Hovenkamp has observed, similar associations of real estate agents, dentists, and even the National Collegiate Athletic Association have tended to collude in ways that disfavor the public interest.[338] Hovenkamp suggests that this problem might be solved by placing Google Search under the governance of a group with more diversity of interest, such as a board composed of "searchers, advertisers, and other market participants who have an independent interest in search quality and product pricing."[339] It is not obvious, however, in what ways regulation by such an outside board would be different from the more traditional remedy of regulation by a specialist public utility regulatory commission, or why it would be better.

---

335. *Id.* at 410–12.

336. Associated Press v. United States, 326 U.S. 1, 4–18 (1945).

337. *Interoperability Remedies*, *supra* note 279, at 19–20.

338. Herbert Hovenkamp, *Fixing Platform Monopoly in the Google Search Case*, PROMARKET (Oct. 6, 2023), https://www.promarket.org/2023/10/06/fixing-platform-monopoly-in-the-google-search-case/.

339. *Id.*

4. Public Utility Regulation

Public utility regulation offers a simpler, more proven approach than the untested proposals described above. The natural monopoly asset can be spun out from its existing owner to a separately-controlled public utility, thus eliminating the incumbent's incentive and ability to favor its own access to the asset over that of competitors.[340] A regulatory commission can ensure that the utility provides nondiscriminatory service, using longstanding "cost-of-service principles" to set rates that compensate the utility for the cost of doing business plus a reasonable rate of return on the utility's investment.[341] The utility may obtain changes to the rate by initiating a rate case before the commission, and may also request pre-approval by the commission of major investment decisions to ensure that the commission will allow the utility to recover the cost of those investments in future rates.[342]

If this approach seems "heavy-handed,"[343] that is because we are comparing it to an illusory "light touch" alternative that, for the reasons explained above, will not work. Any successful interoperability regime is likely to end up looking quite similar to the traditional institution of public utility regulation, which was shaped over decades by the practical realities of mandating nondiscriminatory access to infrastructural goods and services.

I do not mean to deny the well-documented downsides of public utility regulation. But if the scope of this regulation is guided by the updated principles of natural monopoly set out in this Article, we may have less to fear from it than we may imagine, for three reasons. First, critics of public utility regulation often speak from disappointed idealism: they lament that regulation is not effective at reducing price to marginal cost, and therefore not a good approximation of perfect competition. But as we saw in Part I of this Article, perfect competition is neither real nor desirable in most segments of our economy. My proposal is that we reserve public utility regulation for more compelling social purposes. In the Big Tech case, our goal is to prevent private oligopolies and monopolies from exerting too much control over our republic's information ecosystem. As a tool for opening access to the natural

---

340. In this sense, the updated test for natural monopoly I offer in this Article can define the situations in which we should apply vertical separations remedies. Big Tech has extensively discussed these remedies, like interoperability remedies, in the Big Tech Problem literature, but without clarity as to the principles that should guide their application. *See, e.g.*, *Separation of Platforms and Commerce*, *supra* note 7, at 1081 (noting "the challenge of offering a bounded definition of 'dominant platform[s]'" to which separation remedies should be applied).

341. *See* McNish, *supra* note 232, at 628 (collecting citations describing the cost-of-service rate setting process).

342. *Id.*

343. *See, e.g.*, *Equitable Interoperability*, *supra* note 29, at 1015.

monopoly assets from which those oligopolies derive their power, public utility regulation *is* well-suited to its purpose.

Second, the danger that public utility regulation may be "captured" by the regulated entities is real but sometimes overstated. Are public utilities, such as electric companies, truly more politically influential than other unregulated big businesses, such as auto manufacturers or Big Tech companies? The evidence of the last half century or so tends to point in the opposite direction: as Robert Horowitz observed in *The Irony of Regulatory Reform*, during the deregulatory movement, regulators and other publicly interested reformers acted against the will of regulated public utilities, which could not have happened if those firms had truly "captured" the regulatory process.[344]

Third, the impact of regulation on dynamic competition is also real but overstated, particularly when natural monopoly regulation is well-targeted. To repeat: I am not advocating a regulated "search utility" or "social media utility," but only, perhaps, a "web indexing utility" or "content database utility." The potential for innovation and dynamic competition in the provision of these basic, low-differentiation services is less than in consumer-facing product markets, so the opportunity we lose if regulation turns out to inhibit such innovation is correspondingly limited. The innovation we most want is in the downstream markets that depend on access to the natural monopoly asset. Electricity has been provided under a natural monopoly regulatory regime for more than a century, and over that time period, we have seen enormous downstream innovation and competition in the companies and industries that use electricity as an input.[345] Similarly, though its scope may have been overbroad, the AT&T monopoly successfully extended access to high-quality telephone service across the continent, brought numerous valuable inventions into the world, and was regarded for many decades as a symbol of engineering excellence.[346]

CONCLUSION

My goal in this Article has been to offer a regulatory approach for Big Tech that is both more ambitious and better targeted than some of the approaches that have been popular to date. It is more ambitious because I

344. ROBERT BRITT HORWITZ, THE IRONY OF REGULATORY REFORM: THE DEREGULATION OF AMERICAN TELECOMMUNICATIONS 264 (1989).

345. Ironically, the generativity of our natural monopoly electricity system is sometimes celebrated even by writers who argue against the application of similar principles to the internet. *Broadband Debate: A User's Guid*e, *supra* note 253, at 74.

346. THE MASTER SWITCH, *supra* note 222, at 59.

contemplate reaching into major tech enterprises and forcing the structural disintegration of functions (such as web indexing and self-published content databasing) that have to date been an integrated, internal, proprietary component of their overall enterprise. Moreover, I would apply traditional "price and entry" regulation to those disintegrated assets, with all the invasiveness and problematic incentives that entails. But it is better targeted, because I propose to choose the assets subject to this intervention with extreme care and humility, based on a century and a half of theory and experience with natural monopoly. Perhaps counterintuitively, I contend that such natural monopoly regulation can increase competition by providing a principled basis by which new competitors can access critical inputs to downstream innovation. This policy is worth the risk, I think, not for minor efficiency gains, but to solve a compelling social problem: namely, to protect our economy, society, and republic from the dangers of the concentration of power over information in a small number of large firms.