

Traditional Perspectives

DEFINITIONS AND HISTORY

Before trying to reframe debates that have consumed regulatory theory, it is important to outline some fundamentals of regulation. The discussion will center on what regulation is and how it has evolved over time, as well as what its traditional justifications have been. Since many readers will likely have encountered these themes scattered in a variety of other treatments, the discussion will be as brief as possible.

Brilliant commentators have wisely sidestepped the question of specifying the precise meaning of the murky word “regulation.” For example, in Stephen Breyer’s *Regulation and Its Reform*, “no serious effort is made to define ‘regulation.’”¹ As merely a starting point for discussion, however, I do hazard a simple definition: regulation consists of direct public intervention in private contractual arrangements.² What often makes regulation so controversial is that government is the only social institution capable of legitimately coercing people into action. Regulation is thus “the use of this [coercive] power for the purpose of restricting the decisions of economic agents.”³ Typically, commentators have divided regulations somewhat loosely into “economic” regulations, and “social” or “health, safety and environmental” regulations. The former consists of regulating the practices of an entire industry—historically, this has involved “government-imposed restrictions on firm decisions over price, quantity, and entry and exit.”⁴ Canonical examples involve the regulation of railroads, utilities, air transportation, or telecommunications. The latter type of regulation consists of rules designed to minimize the risks to citizens, employees, and consumers. Rather than regulating the behavior of an

industry in toto, social regulations are developed with regard to specific products or activities—whether it be food preparation, pharmaceutical development, workplace safety, environmental pollution, or municipal building codes.

But what about other forms of government intervention? The most obvious, of course, is public finance, where government uses fiscal policy to set social priorities. The difference is that taxes affect behavior only indirectly by creating incentives and disincentives to certain activities, while economic and social regulation consists of direct public intervention.⁵

Whether antitrust policy belongs in a discussion of regulatory policy, however, is a far more complex and controversial subject. Historically, antitrust and regulation have been bifurcated, and remain so in the minds of prominent observers. Stephen Breyer, for example, draws a contrast between “private anticompetitive behavior,”⁶ the domain of antitrust, and “market failures”⁷ that call for regulation:

The antitrust laws seek to create or maintain the *conditions* of a competitive marketplace rather than replicate the *results* of competition or correct for the defects of competitive markets. In doing so, they act negatively, through a few highly general provisions *prohibiting* certain forms of private conduct. They do not affirmatively order firms to behave in specified ways; for the most part, they tell private firms what not to do.⁸

Herbert Hovenkamp notes that as markets “pass out of the realm of strict agency control and into the realm of private, market-based decision making, *antitrust picks up where the regulatory regime leaves off.*”⁹ Paul Joskow warns how “[e]fforts to mix antitrust policies designed to promote competition with regulatory policies that restrict it can cause numerous difficulties.”¹⁰

A new and different, albeit controversial, perspective is to recognize that “antitrust is nothing if not economic regulation,”¹¹ and that the “attempt to draw a sharp demarcation between antitrust and regulatory objectives is a mistake.”¹² As Thomas Moore points out, “The Sherman Act [is] the most encompassing regulatory statute and the second oldest federal regulatory law. . . . Only recently have economists begun to recognize that antitrust laws are regulatory statutes.”¹³ After an extensive discussion of antitrust issues in the health care sector, Peter Hammer and William Sage conclude that the “most important lesson is that competition can and should meet regulation at an interface, not a boundary.”¹⁴

The advantage of bucking the conventional wisdom should hopefully be apparent—a new bridge between these two traditionally distinct areas of policy allows the development of common analytical tools. As William Baumol and Gregory Sidak observe,

This harmony between regulation and antitrust has three important implications. First, the same basic tools of microeconomic analysis can be employed in one as in the other. . . . Second, changes in technology or other circumstances that permit natural monopoly to give way to competition impart continuity to the relationship between economic regulation and antitrust. Third, many of the thorniest problems in antitrust law . . . are fundamentally regulatory in nature, involving issues such as entry or the pricing of intermediate goods sold to competitors. Thus, the economic scholarship on regulation can in many instances enrich antitrust jurisprudence.¹⁵

Put more starkly, “antitrust and regulation strike at the same rocks”¹⁶—one of the central messages in subsequent chapters is to urge policymakers to treat regulation and antitrust holistically.¹⁷ Once the scope of regulation has been framed as economic and social regulation (including antitrust), a brief historical perspective on developments in these areas might be helpful in gaining some sense of the development of American regulatory policy.

The late nineteenth and early twentieth centuries saw the birth of the first regulatory mandates. In the realm of economic regulation, the most notable examples are the mandate of the Interstate Commerce Commission, founded in 1887 to regulate the railroads,¹⁸ and the mandate for the regulation of civil aviation beginning with the Air Commerce Act of 1926.¹⁹ The most prominent illustration with respect to social regulation was the formation of the Bureau of Chemistry in 1908 in response to scandals surrounding adulterated food and drugs.²⁰ The year 1890 saw the passage of the Sherman Act—the foundational statute in antitrust law that seeks to prevent collusion and monopoly; the Clayton and Federal Trade Commission Acts followed in 1914.²¹

Despite these early forays into regulation, the contours of the modern regulatory state were defined only as a reaction to the excesses of the 1920s and the ensuing Great Depression. As Alfred Kahn chronicles, “[The] Great Depression seemed to justify a thorough reexamination of

the microorganization of our economy and a widespread skepticism about unregulated markets. . . . And it was this attitude, as well as the depth of the depression, that led to the cartelization of the entire economy under the National Recovery Act.”²² Roosevelt’s New Deal saw the creation of a number of prominent regulatory agencies, mostly in the realm of economic regulation. Examples fall into two categories. The first consists of those designed to make markets work better, such as the Federal Deposit Insurance Corporation (1933) to protect consumers from unscrupulous banks, and the Securities and Exchange Commission (1935) to rein in Wall Street’s excesses. The second category of interventions sought effectively to manage industrywide cartels—for example, the Federal Communications Commission (1934) to manage the nation’s telephones and airwaves, and the Civil Aeronautics Authority (1938) to set airline rates. The 1930s were a significant decade in American regulatory policy not only due to the sheer volume of agencies created, but also because of the virtually unlimited faith progressive ideology placed in the expertise of these administrative bureaucracies.²³ Stephen Breyer points out this marked shift:

When administrative agencies began to grow in power at the beginning of the twentieth century, the source of their authority was seen to be Congress, and the role of administrative law was to ensure that the agency faithfully carried out its congressional mandate. The agency was a “transmission belt,” applying a congressional statute to changing factual circumstances. . . . *This procedural model broke down with the New Deal.* As congressional delegations were made in broader and broader language, the traditional model could not adequately check the agency’s power.²⁴

Chapter 7 will return to this observation, since avoiding the institutional errors of the New Deal will be central to crafting better regulatory institutions.

With World War II, the Korean War, and the economic expansion of the 1950s and early 1960s, few new regulatory mandates emerged. Antitrust enforcement, for its part, became increasingly strong, especially with respect to merger reviews—even prompting some commentators to speak of the emergence of a newly reinvigorated Sherman Act.²⁵

The cultural revolution and Lyndon Johnson’s plans for a “Great Society” brought the expansion of social regulation in the mid-1960s. Ironically, the pace of new social regulation reached a feverish pitch during

the otherwise conservative Nixon administration in the early 1970s. Kip Viscusi writes:

The decade of the 1970s marked the emergence of almost every major risk or environmental regulation agency. The U.S. Environmental Protection Agency [1970], the National Highway Traffic Safety Administration [1970], the Consumer Product Safety Commission [1973], the Occupational Safety and Health Administration [1972], and the Nuclear Regulatory Commission [1975] all began operation in that decade.²⁶

Much of this social regulation still stands, at least in form; however, as chapter 6 will explore, it has been cut back in practice primarily through misapplication of the principles of economic cost/benefit analysis. As the macroeconomic dislocations of the late 1970s set in, however, much of economic regulation began unraveling.²⁷ Many formerly regulated industries—notably in the transportation and telecommunications sectors—began to undergo dramatic change.

As Alfred Kahn notes, “[B]y 1981 the federal government would have deregulated or substantially deregulated trucking, the airlines, the railroads, and financial markets, and made some progress toward deregulating communications.”²⁸ Two points cannot be overemphasized in this regard. The first is that this movement was, at least as practiced by its leaders, focused primarily on undoing the “cartelization” of the economy wrought by certain aspects of the New Deal. The central idea was to foster competition by “subjecting protectionist regulations that suppress competition to much more critical scrutiny.”²⁹ The second—no doubt due to the sheer force of public intellectuals such as Kahn—is that the deregulation movement benefited from widespread support across the political spectrum for a fight against industry “cartels.” In the context of airline deregulation, for instance, the stunning scope of collaboration produced

an immensely diversified and ultimately irresistible coalition of politicians and private and public interest groups across the country—Senators Edward Kennedy and Howard Cannon, Presidents Ford and Carter, the Consumer Federation of America, Common Cause, Ralph Nader and his various followers, the National Association of Manufacturers, along with such established airlines as United and Frontier, and aggressive would-be entrants like Southwest.³⁰

As industries were being subjected to the vagaries of competition, anti-trust enforcement waned during the 1970s and receded even more dramatically during the 1980s,³¹ with the Reagan administration's appointment of William Baxter to be head of the Department of Justice's Antitrust Division.³²

How to summarize these twists and turns? While ascendant in the early part of the twentieth century, regulation has suffered a gradual decline in respectability over the past thirty years. As Paul Joskow and Nancy Rose remark:

The massive economic disruptions of the 1930s gave rise to a vast array of federal regulations, most of which persisted through the next forty years. The recent wave of federal regulatory reforms arose from the substantial supply shocks and macroeconomic disturbances of the 1970s, which have been characterized as the most severe disruptions since the 1930s. These reforms have dismantled or refigured much of the 1930s federal regulatory apparatus.³³

The wisdom of recent deregulatory fervor is left to future chapters. For now, it is important to understand the justifications for regulation that historically have been used to defend its existence. These arguments serve as one pole of the conventional regulatory debate—the other pole being the neoclassical economic critique of regulation discussed in chapter 2.

JUSTIFICATIONS FOR REGULATION

From an often bewildering array of arguments that permeate the literature, a taxonomy of five classical justifications for regulation emerges—three are economic (natural monopoly, externality, informational imperfection),³⁴ and two are traditionally considered noneconomic (redistribution and paternalism). As defined above, regulation is government intervention into private contractual arrangements: economic justifications of regulation, then, must necessarily show “why markets may fail, i.e. why mutual gains are not necessarily fully exploited by decentralized decision-makers.”³⁵

The existence of a natural monopoly has, historically, provided the quintessential economic justification for regulation.³⁶ In these markets, cost per unit decreases as output level increases,³⁷ making it most efficient for a single large firm to provide service.³⁸ Such a situation typically evolves in industries where “there is a large fixed-cost component to

cost”³⁹—prototypical examples are energy or telecommunications utilities. The reason why an unregulated natural monopoly is economically wasteful is that the monopolist “will restrict output to attain a higher price and higher profits than is possible under free competition between many small sellers.”⁴⁰ This economic logic is the reason why basic infrastructure industries were the first to be regulated in the late nineteenth century. As Roger Noll points out,

In the United States, the first examples of regulatory programs were justified on the basis of natural monopoly: a specific good or service (in this case, grain elevators, water supply, and railroads) could be produced at lowest cost only if supplied by a single firm, but this would give rise to monopolistic abuse and dead-weight loss in an unregulated market.⁴¹

Regulators controlled entry and exit into an industry to prevent abuse of scale economies, set rates to protect consumers, and even dictated what products and services the monopolist could offer.⁴² In many ways, natural monopolies provided the temptation to overreach and engage in “command and control” regulation in the belief that regulators could manage all aspects of an industry. This temptation will be revisited in chapter 7.

The second economic justification for regulation is that of “externality.” Simply put, externalities arise “when economic agents impose costs on, or deliver benefits to, others who are not parties to their transaction.”⁴³ In other words, economic actors are affecting the welfare of society, yet are not made to internalize this effect: these differences “between true social costs and unregulated price are ‘spillover’ costs (or benefits).”⁴⁴ The result is a mismatch between private and social welfare.⁴⁵ Most prominently, externalities have provided justification for environmental regulation;⁴⁶ after all, as one book on welfare economics warns, “polluters do not have to pay for the losses they inflict on others. This means that the resource allocation provided by the market may not be efficient.”⁴⁷

Positive externalities, for their part, are intricately related to the frequently confused notion of “public goods.” The link is to recognize that private market transactions typically do not have the incentive to produce goods that have positive externalities—these public goods include parks, national defense, roads and highways, and even broadband infrastructure.⁴⁸ After all, “since consumers cannot be excluded from consuming a pure public good it is very difficult to provide such goods

via the market. *Who is willing to pay not only for their consumption, but also for everyone else's consumption?*"⁴⁹ The result is a need for government intervention, as James Q. Wilson notes:

[G]overnment (that is, an institution wielding coercive powers) must compel us to pay for national defense or clean air because the market will not supply these things. The reason is simple: Since everybody will benefit from national defense or clean air if anyone in particular benefits, no one has any incentive to pay for these things voluntarily. We all would be free riders because none of us feasibly could be excluded from these benefits if he or she refused to pay.⁵⁰

The third economic rationale for regulation is to remedy informational imperfections among buyers and sellers, the consumers and producers, in private markets. As Stephen Breyer argues:

Markets for information may on occasion not function well for several reasons. First, the incentives to produce and to disseminate information may be skewed. . . . Thus, those in the best position to produce the information may not do so, or they may hesitate to disseminate it, for fear that the benefits will go not to themselves but only to others. . . . Second, one of the parties to a transaction may seek deliberately to mislead the other, by conveying false information or by omitting key facts. . . . Third, even after locating potentially competing sellers, the buyer may not be able to evaluate the characteristics of the products or services they offer.⁵¹

The presence of these informational imperfections, or asymmetries, has been invoked to justify much social regulation that tries to protect consumers and workers against health, safety, and environmental risks. Roger Noll notes that imperfect information

has been the rationale for regulating consumer products and workplaces, beginning with the Pure Food and Drug Act of 1906. Complex, costly information can lead to poorly informed and sometimes potentially hazardous decisions about goods, services, and jobs. It can also lead suppliers to provide either too much or too little quality, and industries to adopt inefficient technical compatibility standards. In principle, regulation can provide two types of efficiency gains. First, by increasing the supply of information, it can reduce

uncertainties about the consequences of market decisions, thereby causing the market to make a better match between suppliers and demanders. Second, by setting minimum standards, it can protect uninformed participants against bad outcomes, including a “market for lemons” equilibrium in which quality is supplied at inefficiently low levels.⁵²

Needless to say, the economic justifications for regulation are not mutually exclusive. In particular, the externality and informational arguments have frequently been combined in the context of social regulation. As Kip Viscusi cogently argues,

The chief impetus for the health, safety, and environmental regulations is twofold. First, substantial externalities often result from economic behavior. The operation of businesses often generates air pollution, water pollution, and toxic waste. Individual consumption decisions are also the source of externalities, as the fuel we burn in our cars gives rise to air pollution. Informational issues also play a salient role. Because of the peculiar nature of information as an economic commodity, it is more efficient for the government to be the producer of much information and to disseminate the information broadly. Individual firms, for example, will not have the same kind of incentives to do scientific research unless they can reap the benefits of the information. As a result, it is largely through the efforts of government agencies that society has funded research into the implications of various kinds of hazards so that we can form an assessment of their consequences and determine the degree to which they should be regulated.⁵³

Beyond the three interwoven economic justifications for regulation, there are, as mentioned, two often-overlooked traditionally noneconomic justifications that have little to do with price or output per se: redistribution and paternalism. Here, moral and political rather than economic argument often reigns. As one commentator notes, “[F]rom a political view, perhaps the most significant feature of regulation is that it redistributes income.”⁵⁴ Thus, arguments for regulation that revolve around controlling the “rents” or “excess profits” producers enjoy are not efficiency based; rather, the “justification rests upon the desirability of transferring income—from producers to consumers.”⁵⁵ Much to the chagrin of many

neoclassical economists, distributional goals are especially stark in economic regulation. Examples include “universal service” in telecommunications, which funds telephone services to low-income consumers,⁵⁶ the “duty to serve” of electrical utilities,⁵⁷ and “diversity” and “localism” in media.⁵⁸ Rationales behind these mandates “argue for regulation for reasons related not to more efficient use of the world’s resources, but to a fairer income distribution.”⁵⁹ In a closely analogous manner, congressional intent behind the antitrust laws reveals a concern for consumers and small businesses, not for overall economic efficiency.⁶⁰

A related rationale for regulation is paternalism—that government can help shape preferences when individuals act in self-destructive ways:

It is sometimes argued that individuals don’t necessarily behave in ways that are consistent with what is in their own best interest. Not all of us voluntarily use seat belts or complete elementary education, although we should. This argument for government intervention is called *paternalism*. It expresses a particular view concerning the extent to which one should respect consumer preferences, and has nothing to do per se with the presence or absence of market failures.⁶¹

It is important to dwell on the traditionally noneconomic rationales for regulation, precisely because they are either ignored or disparaged in the mainstream law and economics literature. As Gregory Sidak and Daniel Spulber observe, “Economists may posture as purists and assert that regulators are misguided to pursue any goal other than economic efficiency. However correct that position may be as a matter of theory, it does not take the institutional setting of regulation as it really is.”⁶²

Is there a common framework for these five seemingly disparate justifications for regulation? All of the arguments made in this section can be combined under one rubric: government regulation is designed to protect the “public interest.” More precisely, economists have labeled the “public interest” view of regulation as “normative analysis as a positive theory” (NPT).⁶³ NPT combines a positive description of inefficient or inequitable markets with the normative goal of rectifying these problems through government intervention.⁶⁴ As Viscusi points out, NPT

uses normative analysis to generate a positive theory by saying that regulation is supplied in response to the public’s demand for the correction of a market failure or for the correction of highly inequi-

table practices (for example, price discrimination, or firms' receiving windfall profits as a result of some change in industry conditions).⁶⁵

The blending of positive and normative arguments is part of a long intellectual tradition, dating back at least as far as David Hume's observation that "is" and "ought" often blend in moral reasoning.⁶⁶ As one prominent legal scholar observes, "We have no methodology to move directly from the discourses we perceive as descriptive, such as natural or social science, to decisions about the way to organize our society and the kinds of laws we should establish to effect that organization."⁶⁷ NPT at least reflects this intellectual realism. Yet its underlying premises have suffered a massive, perhaps even fatal, attack from a group of intellectuals cleverly portraying themselves as the "law and economics" vanguard.⁶⁸