

23. Licit and Illicit Responses to Regulation

LEE BENHAM

1. INTRODUCTION

New regulation can elicit a great variety of responses from individuals, firms, interest groups, and bureaucracies. This chapter examines a range of common legal and illegal behaviors that arise as responses to new regulations. It also compares the approaches that new institutional economics and neoclassical economics use to study these responses. The motivation for introducing new regulation is generally to influence behavior: to promote or restrict competition, to redistribute income, to increase or reduce barriers to entry, to increase or reduce spillovers, and so on. However, regulation often influences behavior in ways that differ from the initially stated rationale. This chapter focuses on the consequences of regulation, on this wide range of responses, rather than on the rationale offered for introducing the regulation.

In economics, the standard literature on regulation emphasizes two dimensions of response—money price and quantity—and neglects other dimensions. This limited focus reflects the small number of variables emphasized in the standard theory, a desire for parsimony, and limited data. However, responses across many dimensions are possible. There is no *a priori* ground for believing that the standard price-and-quantity responses or the income effects are always the most significant ones. We need to discover which consequences are big and which are small, to learn how various effects play out over time, and to understand the conditions under which a particular type of response is likely to occur. On occasion, a single small regulation may cascade into a large regulatory system. At other times, an apparently Draconian regulation may be greatly weakened in its effects by inventive adaptations along various margins.

The range of possible licit and illicit responses to regulation shown in Table 1 illustrates some of the approaches taken by new institutional economics to the study of regulation. Although far from exhaustive, this set is much greater than the set usually examined in the standard literature on regulation.

Table 1. Categories of responses to regulation

<i>Licit responses</i>	
<i>Substitute</i>	<i>for</i>
Other goods	Regulated good
Other attributes of good	Regulated attributes
Amenities	Profits in excess of regulated maximum
Barter and other arrangements	Money
Vertical integration	Market exchange
Household production	Market production
Personalized exchange	Impersonal exchange
<i>Alter</i>	
Governance and contractual relationships	
Organization of the market	
Interest groups and their goals	
Other formal regulations	
Informal norms	
<i>Illicit Responses</i>	
<i>Vary the extent of</i>	
Underground economy	
Private coercion and extralegal organizations	
Discrimination	
Corruption	

2. LICIT RESPONSES TO REGULATION

Substitution of Other Goods

If a regulation raises the price of a given good, consumers will tend to consume less of that good and to substitute other goods. The standard neoclassical model focuses principally on this effect.

At a given time and place, the set of goods available in the marketplace is merely a subset of all goods that are potentially available. The set existing in the marketplace is established as the outcome of an equilibrium process that depends on competition, production possibilities, income, and the formal and informal rules of the game. Changes in regulation can affect these factors and hence change the available set of goods. This can alter the range of substitutes available to consumers.

Substitution of Attributes of Goods

Regulation also affects the equilibrium attributes of goods. Even “simple” goods have many attributes beyond their money price, such as size, color, quality, reliability, warranty, availability of credit, associated service, location, and waiting

time.¹ All of these can vary with market conditions and regulations. The standard economic model assumes that money price and quantity are the two dimensions along which adjustments take place between buyers and sellers, and that these adjustments continue until net gains are zero for the marginal unit exchanged. By these adjustments, sellers and buyers minimize the deadweight loss. As Yoram Barzel (1997) has discussed, the same logic applies to the non-money attributes of goods: buyers and sellers jointly adjust these other attributes until the net gains are zero for marginal changes there. When a regulation alters one attribute (or a few—it can never control all of them),² it is in the mutual interest of buyers and sellers to minimize the deadweight loss by varying other attributes.

Depending on the substitution possibilities, therefore, regulation of money prices can lead not only to shortages or surpluses and welfare losses, but also to changes in other attributes. As an example, consider the price ceilings on gasoline in the U.S. in effect from 1971 to 1974. Faced with regulated prices set below the market-clearing level, service stations reduced their hours of business to the minimum needed to sell their allocation, lowered their quality of service, and offered lower-octane fuel. Customers waited longer in line to obtain gasoline and purchased new vehicles designed with larger gas tanks. Buyers and sellers jointly sought the lowest cost adaptation to the regulation.

Hong Kong rent controls provide another illustration. The first regulations, imposed in 1921, prevented landlords from raising rents except when they demolished an existing building and replaced it with a new one. In such cases, existing tenants received no compensation. A building craze ensued. By 1923, some landlords were even replacing buildings that were only two years old, and many former tenants were sleeping in the streets.³ The building owners captured some of the returns from attribute substitution (a building just constructed versus a building in use), but many potential gains were dissipated by excessive construction. Eventually, in 1955, under another rent control regulatory system, the government introduced an enforced-compensation scheme

¹Economic theory generally treats goods as purely homogeneous: if two items differ in their attributes, they are treated as being different goods. By this view, a liter of gasoline that requires 30 minutes of waiting time to obtain is a different good from a liter of gasoline that requires zero minutes of waiting time. This is analytically useful, but in empirical analysis this problem is not so easily addressed.

²Price controls typically involve attributes that are relatively straightforward for the regulators to measure and monitor—number of units, weight, and money price, rather than aspects that are more costly to measure—like quality, reliability, and associated service. Similarly, empirical work tends to focus on attributes that can be measured at low cost to the investigator. Studies of gasoline regulation typically focus on money price, volume, and octane rating (for which data are readily available) and not on quality of service and waiting time. However, real world responses to regulation may also involve attributes that are very costly for regulators and scholars to measure.

³Steven Cheung (1975). Even when rent controls carried over into new buildings, the new tenants would give the owners “key money” for access to the apartments. These side payments could not easily be regulated because they could be disguised as the sale of other items such as old chairs. In principle, key money could be paid on a continuing basis to keep a landlord from tearing down a building, but this was illegal, and it also involved high costs of collective action, since most building had multiple tenants.

for displaced tenants. This new scheme clarified ownership rights and greatly reduced the transaction costs of negotiations between landlords and tenants. It better aligned incentives for recognizing costs borne by displaced tenants and costs for new construction. Steven Cheung argues that the Hong Kong rent control system at that time was highly efficient compared to rent control systems elsewhere in the world. He argues further that Hong Kong's continuing growth through the twentieth century would not have been possible without relatively low transaction costs between tenants and landlords.⁴

Substitution of Amenities

Certain regulations, such as ceilings on money profits in firms, lower the cost to the firms' decision-makers of consuming amenities. Whenever the money profits of a regulated firm are potentially greater than the maximum permitted, this reduces the opportunity cost to the firm of having elegant company dining rooms, chauffeured cars, congenial colleagues, good relations with the unions, nepotism, and the quiet life (Armen Alchian and Reuben Kessel 1962). During the years when U. S. banking regulations severely constrained competition among banks, imposing fixed interest rates, restricted entry, and limited branch banking, the top executives in that industry should have led less stressful lives than their counterparts in more competitive industries. A study of executives' life expectancy during that time period found that bankers indeed had a longer average life expectancy than did executives in other industries. This greater longevity is consistent with the view that these regulations reduce the costs to decision-makers of choosing a desirable lifestyle. (Gili Yen and Lee Benham 1986)

Substitution of Barter and other Means

When regulation lowers the transaction costs of using barter instead of money, barter will increase. Trade restrictions and foreign exchange controls can have a major impact. In recent years, approximately 10% of world trade has been conducted through barter.⁵ Historically, barter on a wide scale has been a common occurrence. Consider the case of Germany in the period around World War II. In 1936, before the war, the German government froze prices. At the end of the war, the occupying military governments retained these controls. In 1947 the money in circulation was ten times as great as in 1936, and real income had fallen by half. Official prices did not reflect existing scarcity, and the black

⁴Cheung (1979). Over time the tenant-landlord problem also diminished because builders increasingly sold apartments as condominiums rather than rent them out. This movement toward owner-occupants rather than tenants is a form of vertical integration which addresses some problems of incentive alignment. It depends on the existence of capital markets with low transaction costs for prospective condominium owners.

⁵Even in open market conditions, barter is still sometimes the lowest-cost way to trade. Barter exchange among U.S. corporations is growing, partly because computers allow low-cost measurement and tracking of barter arrangements. See Akbar Marvasti and David Smyth (1998).

market was confined to a narrow range of consumer goods. Barter was illegal, but in order to obtain raw materials, firms developed a complicated chain of barter arrangements. Even under the direct control of the British military government, Volkswagen utilized 5% of its production for barter arrangements. Legal wage levels were set too low to secure a regular supply of employees, so employers compensated employees with goods in kind, which could then be used to barter for food (Wendy Carlin 1989). Against the devastation of the war and the harsh regulations of the period, barter and other adaptations kept real income and output from falling further.

In addition to barter, many other substitutes for money arise where they are cost-effective. Radford gives an instructive example in his study of the economics of a prisoner-of-war camp (R. A. Radford 1945). Money was very limited in the camp, and barter began. This was then superseded by the use of a money substitute as the medium of exchange—cigarettes, which were reasonably homogeneous, divisible, and durable.

Substitution of Vertical Integration

Regulations such as price controls and cartel pricing can alter the relative advantages of vertical integration versus market exchange. This was strikingly illustrated in the U.S. during and immediately after World War II, when the imposition of price controls led to a substantial increase in vertical mergers. In the late nineteenth century, the Rhenish-Westphalian Coal Cartel increased coal prices substantially. Many electricity and gas utilities, railroads, and even municipalities then acquired their own coal mines (Archibald H. Stockder 1932).

Regulations that affect the existence or costs of particular markets, such as futures markets, also affect the incentives for vertical integration. A futures market, among many roles, serves as a synthetic storage mechanism and an alternative to vertical integration into storage. When the futures market in oil began in 1983 in the U.S., this lowered the cost to firms of using the market to obtain oil in the future as compared with storing the oil themselves. This was followed by a reduction in vertical integration in the oil industry and by changes in the terms of contracting (Michael Sykuta 1994 and 1996).⁶

Substitution of Household Production

Households produce goods and services both for their own consumption and for trade with the external market (Yoram Ben-Porath 1980). As Robert Pollak describes it (1985, pp. 605–606),

“The transaction cost approach views marriage as a ‘governance structure,’ emphasizes the role of ‘bargaining’ within families, and draws attention to the advantages and disadvantages of family organization in terms of incentives and monitoring,

⁶The literature in economics and finance traditionally viewed futures markets as hedging mechanisms, not as substitutes for vertical integration. See Marshall (1920) and Keynes (1930).

and to the special roles of ‘altruism’ and ‘family loyalty’. It also recognizes the disadvantages of family governance: conflict spillover, the toleration of inefficient personnel, inappropriate ability match, and inability to realize economics of scale. If activities are assigned to institutions in an efficient or cost-minimizing fashion, the balance of these advantages and disadvantages plays a major role in determining which activities are carried out within families and which are performed by firms, nonprofit institutions, or the state.”

Levels of transaction costs within the family and within the external market will affect the extent and type of household production. Regulations that increase the costs of impersonal exchange in the market will increase the importance of household production and other forms of personalized exchange. In the Soviet Union, when the market system was severely repressed, families vertically integrated into food production for themselves by working on their small garden plots in the countryside. In many settings, land regulations limit individuals’ ability to obtain property rights that are clear and protected. This leads to difficulties of finance and contracting. Houses built in such settings are often constructed by the hands of the household members, with bricks purchased when cash is available (Hernando de Soto 1989).

Substitution of Personalized Exchange

One way to characterize regulatory regimes is the extent to which there is impersonal exchange, personalized exchange, or little exchange at all. The extent of impersonal exchange affects the size of the market and the degree of specialization, as Adam Smith (1776) observed. Indeed, Douglass North (1991) has described the process of economic development as a movement from personal to impersonal exchange.

Impersonal exchange requires low costs of measurement and enforcement. Adam Smith’s invisible hand depends upon a regulatory regime that encourages impersonal exchange by keeping transaction costs low. By altering such costs, regulation affects the extent of impersonal exchange. Some regulations like standardization of weights and measures can lower measurement and enforcement costs and thus promote impersonal exchange. Many other kinds of regulations raise transaction costs and thereby promote more personalized exchange. When regulatory regimes are arbitrary,⁷ impose excessive entry costs, establish price controls, or generally cause high transaction costs, then personalized exchange is likely to be widespread. For example, in late twentieth-century Egypt, markets for many goods were heavily regulated. Shortages arose and it was often difficult to find goods in the impersonal market. Egyptians adapted in part by developing their own reciprocity networks of personalized exchange. These informal personalized networks created more opportunities to locate rationed goods or to find jobs (Diane Singerman 1995).

⁷Trade that involves political favors usually involves personalized exchange.

Changes in Governance and Contractual Relationships

Firms and households shape their governance structure—the forms of contracting, ownership, and decision-making—in part by efforts to economize on transaction costs. Since regulations affect transaction costs, changes in regulations can affect governance structures. This argument holds a central place in the work of Oliver Williamson (1985, Williamson and Masten 1995) and in the associated literature on transaction cost economics.

To illustrate, consider the impact of unit banking regulations in the U.S. Historically some U.S. states prohibited bank takeovers by other banks, while other states did not. In states that prohibited takeovers, the banks were less efficient and had lower profits. Why? A competitive constraint on inefficiency had been removed. Where takeovers were not possible, managers were subject to less outside scrutiny, so more inefficient practices could and did arise. This was somewhat mitigated, however, by changes in bank ownership structure. If bank ownership is more concentrated, some individual shareholders are more likely to monitor the managers. This is because the benefits from monitoring, being divided among a smaller, concentrated set of shareholders, are more likely to offset monitoring costs for at least some shareholders. In the U.S. case, in the states that prohibited takeovers, ownership concentration and management ownership were indeed higher. These mechanisms reduced the inefficiencies associated with the restrictions on takeovers, but they were not perfect substitutes for a takeover market (Mary Schranz 1993).

Changes in Organization of the Market

In many markets, regulatory standards exist for advertising, disclosure, and measurement. These regulations affect the types of information and goods produced and can give rise to dramatically different configurations of economic organization—sometimes improving efficiency and sometimes not. In the case of the market for eyeglasses, regulations in many U.S. states historically restricted severely the types of advertising that could be provided to consumers. This greatly limited the ability of the large firms, which were the low-cost providers, to compete. These regulations were also associated with higher prices of eyeglasses to consumers, lower frequency of purchase, and more adverse effects for those with less education (Lee Benham 1972, Lee Benham and Alexandra Benham 1975).

Regulations that lower the costs of measurement, such as standardization of weights and measures, can increase enforceability and credibility, thereby enhancing the efficiency of markets. Regulations of quality standards for food can increase the credibility of brand names. Together these appear to have assisted the development of the market for manufactured foods. Food-manufacturing firms in the U.S. at the beginning of the twentieth century recognized this, and they themselves supported the introduction of quality-control regulations (Marc Law 2003a and 2003b).

Note that even if a regulatory environment lowers the costs of measurement and enforcement, this does not mean that the total resources in the economy devoted to transaction costs will necessarily decline. The transaction costs per transaction may fall, but the total number of transactions may increase dramatically, so that total transaction costs rise. John Wallis and Douglass North (1986) have evidence that bear on this issue. Between 1870 and 1970 in the United States, specialization and per capita income grew enormously. Simultaneously, the transaction sector—those activities such as accounting and law which involve measurement and enforcement—increased from 25% to 45% of GNP.

3. ILLICIT RESPONSES TO REGULATION

Development of Underground Economic Activity

The underground economy, the “informal economy,” includes (a) legal activities conducted outside the formal legal system (manufacture of legal goods without a permit, providing services like house repairs without reporting income to tax authorities) and (b) illegal activities (illegal drugs, stolen goods, prostitution, gambling). Informal economies play a significant role throughout the world; their relative importance varies by country.⁸ Available evidence suggests that variation in regulatory climate is a major reason for variation in the size of the underground economy.

Of particular relevance here are barriers to entry, perhaps the most intensively studied form of all regulation. Adam Smith gave them a central place in *The Wealth of Nations* (1776). Hernando de Soto (1989) documented startlingly high regulatory barriers to small-business entry in Peru in the 1980's.⁹ A great proportion of business in Peru was transacted informally. Djankov et al. (2002) studied of regulations governing business entry in 85 countries and found high official costs of entry in most countries. Djankov (2003) also found that where regulation of entry is more extreme, the unofficial economy is larger, corruption is higher, and the quality of public service is not better.

Development of Discrimination

Regulations can alter the incentives to discriminate and thereby influence the extent of discrimination. Through rewards and sanctions they can directly alter the price of discriminating. They can also work indirectly. For example, if the money profits of a regulated firm are restricted below the level achievable, this

⁸In the year 2000, the share of the economy that was informal averaged 41% in developing countries, 38% in transition countries, and 18% in OECD countries. Differences within regions were substantial: Bolivia 67% vs. Chile 20%; Zimbabwe 59% vs. South Africa 28%; Thailand 53% vs. Japan 11%; Greece 29% vs. Switzerland 9%. See Friedrich Schneider (2002).

⁹In an experiment, de Soto demonstrated that a person seeking legal permission to open a small clothing factory on the outskirts of Lima in 1983 needed to spend 289 days of full-time effort and pay an irreducible minimum of two bribes to obtain the necessary permits from the bureaucracy.

will lower the cost to the firm of employing workers who are less productive but who have preferred characteristics along other dimensions. In other words, it will lower the cost to the firm of engaging in discrimination.

The regulation of U. S. physicians in the early part of the twentieth century provides an illustration. At that time, medical licensure came under the monopoly control of state medical associations. This strengthened the medical associations' position as an interest group, and they achieved a dramatic reduction in the number of medical schools. As a result, physicians' earnings increased, and applications to surviving medical schools increased. The relative numbers of women and blacks admitted to medical schools then declined sharply, as regulatory-induced entry barriers, waiting lines of applicants, and property rights and organization within medical schools lowered the cost of discrimination to those making the admission decisions (Kessel 1970).

Development of Private Coercion

Regulations that raise the cost of engaging in voluntary exchange will affect the extent to which private coercion is used, individually and through illegal organizations. Regulations such as prohibitions on alcohol, drugs, and gambling increase the potential gains to dealing in these arenas and thereby facilitate the rise of illegal organizations like the Mafia.

Regulations that increase uncertainty concerning ownership of property rights can lead to environmental problems like land invasions, violence, and deforestation. In Brazil, the conflicted regulatory environment concerning property rights has led to all of these (Lee Alston 1999; Alston et al. 2000). In Italy in the nineteenth century, given the weakness of the relevant formal institutions, changes in property rights to land were followed by a great increase in the strength of the Mafia, which had a comparative advantage in contract enforcement (Diego Gambetta 1993).

Development of Corruption

Corruption is the use of public office for private gains in carrying out a public task. Regulations can significantly affect the benefits and costs of engaging in corrupt practices. In settings where regulations are excessive or arbitrary, where civil servants are poorly paid, where the level of political competition is low, where transparency is lacking, where civil servants have discretion over enforcement, where discretionary permits are highly valuable, and/or where the likelihood of exposure is low—in these settings corruption is likely to be extensive.

The number of possible corrupt practices is very large. The organization of corruption and the property rights to associated payoffs vary considerably by country (Andrei Shleifer and Robert Vishny 1993). Where the property rights to bribes are clearly held and enforced, a given level of corruption may produce less inefficiency. If the highest-ranking politicians and political parties are disciplined enough to refrain from seeking additional payoffs after initial lump

sums have been paid, corruption on the margin may be circumscribed. In South Korea, for example, political parties are highly disciplined, and bribery has been highly centralized. Many payoffs are lump-sum contributions by major business leaders to presidential campaigns. This arrangement does not tax economic activity at the margin. If the government is weak and fragmented, however, local government regulators often engage in decentralized looting. This undermines credible commitments in both public and private sectors and can be particularly damaging to economic performance (Pranab Bardhan 1997).

Regulators are often active in this process. They may intentionally introduce regulations to create further opportunities for corrupt exchanges. If corruption becomes common in one sector (say, because of high import duties or restrictions on gambling), the cost of spreading corruption to other sectors falls. This in turn lowers the returns to normal entrepreneurial activity, which further slows down economically productive activity, which induces more people to engage in corrupt practices, and so on (Kevin Murphy et al. 1993). Furthermore, it may be easier to introduce corrupt practices than to reduce them: long periods of time may be needed to move from a corrupt to a non-corrupt equilibrium.

4. HISTORY, PATH DEPENDENCE, AND INTEREST GROUPS

The discussion so far has focused on the responses to regulation made primarily by individual actors, acting on considerations of supply and demand. Let us turn now to the role of history and path dependence. In most of the standard literature on regulation, these have been ignored. However, historical experience is fundamental to the formation and evolution of formal and informal rules, as well as to the consequences of any new formal regulation. Furthermore, one regulation often leads to others, and the historical reach can be long. There is path dependence in regulation, substantially rooted in the political process and in historical experience.

Oliver Williamson's classification of the different levels of institutional constraints, shown in Table 2, provides a perspective on the different time frames involved. Neoclassical economics has focused almost entirely on the bottom level, resource allocation and employment, where adaptation to regulation is restricted to changes in prices and quantities and where adaptation is continuous. However, countries have differing historical experiences, norms and traditions. In Williamson's terminology, "embeddedness" differs. The specifics of Table 2 can be disputed; for example, norms and religions may change more frequently than every 100–1000 years, but this perspective is valuable in emphasizing the varying time spans over which institutional changes can take place. At any point in time, the formal rules, informal norms, and their enforcement characteristics vary substantially across countries. The impact of any new formal regulation—and responses to it—will be affected by these elements and by their historical patterns within a country.

Depending on context and time period, therefore, a particular type of regulation can give rise to very different responses in different settings. The direct

Table 2. Williamson: economics of institutions¹⁰

<i>Level of Embeddedness</i>
Informal institutions, traditions, norms, religion
Frequency of change: 100–1000 years
<i>Level of Institutional Environment</i>
Formal rules of the game—especially property (polity, judiciary, bureaucracy)
Frequency of change: 10–100 years
<i>Level of Governance</i>
Play of the game—especially contract (aligning governance structures with transactions)
Frequency of change: 1–10 years
<i>Level of Resource Allocation and Employment</i>
Prices and quantities, incentive alignment
Frequency of change: continuous

allocative effects of a regulation—measured at Williamson’s levels of resource allocation and governance—are often only a small part of its total impact. Small and seemingly unimportant details of regulations can lead to large downstream effects.¹¹ Or apparently far-reaching regulations may turn out to be inconsequential because of innovative adaptations over time. While standard microeconomic analysis of regulation invokes the notion of the long run to some extent, in the perspective of new institutional economics the long run is usually longer, more varied, and more significant. The following case studies illustrate how regulations, constrained by path dependence, can generate long-run responses by interest groups with further policy consequences.¹²

Broadcasting in Britain

Small initial regulatory choices can lead to large downstream effects. Ronald Coase found this in his detailed investigation of the origins of radio broadcasting in Britain (Coase 1947). Initially, the transmission of sound by radio was regarded as a new means for sending telephone messages. Because of this, regulation of broadcasting was assigned to the organization that had jurisdiction over telephones: the Post Office. To avoid the problem of broadcast interference and to avoid having to select from among the many firms who would seek the

¹⁰Table from Oliver Williamson (1999), p. 11.

¹¹The impact of any regulatory change will depend in part on the informal rules. If a regulatory change is greatly at variance with the preferences and interests of a particular group, their informal norms can evolve into “opposition norms.” These opposition norms can have highly negative consequences for performance. See Victor Nee (1998).

¹²Three major interest group theories of the regulatory process are:

- a. The public interest paradigm: government is a benevolent guardian.
- b. The special interest “capture” theory: regulations come from the demanders of regulation and not from the suppliers or the government. At the limit, this view holds that special interests are responsible for the origins of the regulation.
- c. The public choice view: special interest groups are the demanders of regulation, and politicians and bureaucrats are the suppliers, all working in their own self-interest.

valuable radio spectrum, the Post Office favored having a single broadcasting company.¹³ Strategies that could have been used to allocate the radio spectrum, including assigning property rights to the spectrum and selling them off, were not seriously considered.¹⁴ The main advocates for radio broadcasting at that time were the radio manufacturers, who sought their revenues from the sale of radios rather than from broadcasting. Newspapers did not want competition for their advertising revenue and therefore opposed commercial forms of radio broadcasting. There was thus little opposition to the Post Office's proposal to create only one broadcasting company and to finance its programming and transmission from taxes on radio sales. There followed a broadcast monopoly that limited British citizens to BBC programming for many decades. Non-BBC programming could not even be transmitted by wire (although that involved no problem of externalities). When television was eventually introduced, the BBC monopoly was extended to cover television.

In the years just before World War II, the BBC gave the major political parties control over access to the airwaves for political speeches. The political parties then allocated their shares of airtime to politicians who endorsed their standard party positions. When Winston Churchill sought to criticize the government's appeasement policy towards Hitler, he was unable to get permission to speak on the BBC. Given the rules of the game, this result is not surprising. An apparently small regulatory choice driven by specific historic circumstances had these profoundly important later consequences.

U. S. Sugar Production

Sugar production in the U. S. offers another example of the downstream impact of initial regulatory decisions. Sugar is the only U.S. agricultural good now excluded from global free trade. In 2002 raw sugar prices in the U.S. were three times as high as the world spot price. All 11,800 U.S. sugar farmers receive subsidies. As a result of these subsidies, the Everglades region in Florida has been planted with sugar cane, creating significant ecological problems. Sugar-processing and sugar-using industries have moved out of the U.S. because of the high domestic sugar prices. The total costs of this program are vast. How did all this happen?

Anne Krueger's classic study (1990) shows how the complexity and the cost of current sugar regulation can be understood only by an appreciation of past policies. Through most of the nineteenth century, 90% of the sugar consumed in the

¹³C. A. Lewis, deputy director of the British Broadcasting Company, stated in 1923, "The chaotic state of affairs in America, where a large number of stations are transmitting on a narrow band of wavelengths and no form of control exists, was an object lesson in what *not* to do, and consequently the control was put into one company's hands." Lewis (1924), quoted in Coase (1947), p. 208, fn. 2.

¹⁴"Only many decades later were regulations introduced to define some rights to some spectrum such that sales could take place. The common view at the time among experts was that the regulation came as a necessary result of reckless competition among broadcasters, which retarded the orderly development of radio and subjected the listeners to intolerable strain." Coase (1959), p. 13.

U.S. was imported. Tariffs on sugar were originally imposed as a revenue source. Then in the late 1800's the U. S. Department of Agriculture took initiatives to encourage domestic sugar growing and processing. The sugar industry developed because of this government-initiated program, rather than the reverse.¹⁵

Several historical accidents appear to have had a major impact on the subsequent evolution of the program, including idiosyncratic committee assignments in Congress and a focus on foreign relations issues. Once the very complex sugar program was in place, a network of program specialists arose in government and industry and became independently influential. Various interest groups were involved over time. A long-standing—and unlikely—coalition of sugar producers, processors, and users persisted until 1978. At that time a new corn-based substitute for sugar began to take market share. Corn producers benefited from having high prices for sugar and they therefore supported a sugar quota system, opposing plans that would have permitted the U.S. sugar price to fall to the world level.

The high domestic price of sugar has led to many market responses: the use of sugar substitutes, the importing of products such as cake mixes with high sugar content, and the subsequent ban on imported cake mixes. Diplomatic complications have been associated with the allocation of sugar import quotas across countries. Although this program has been very costly to consumers and to sugar-using industries, there appear to be few long-term rents earned by domestic sugar producers and processors. Most of the subsidies are dissipated in inducing production where it would not otherwise take place.

Examples from many other industries could be provided. The regulation and deregulation of natural gas, standards for automobile emissions, pharmaceutical regulation, medical licensure, and airline regulation all show a history of interest group politics in which new regulations frequently create new interest groups.¹⁶

The Old Believers in Russia

The previous examples have shown that small initial regulations can lead step by step to major downstream consequences. On the other hand, Draconian regulations are sometimes followed by surprisingly resilient responses over long time periods. Consider the case of the Old Believers in Russia.

Starting in the seventeenth century, the Old Believers were heavily persecuted for refusing to conform to the new religious doctrines of the Russian Orthodox Church. During the period of most intense persecution, the Old Believers endured removal of all their legal rights, forced exile, no standing in court, and even mass executions. Although the intensity of persecution diminished over time, they remained for centuries without full legal status. However, they survived and

¹⁵In 1934, a quota system that in fact benefited the sugar producers was imposed *over their objections*. They learned from experience, and three years later they began to support the quota system.

¹⁶Roger Noll and Bruce Owen (1983). Most studies of collective action have emphasized the opposite direction of analysis: interest groups' efforts to create regulations serving their interest. See Mancur Olson (1982). Mary M. Shirley (2002) has examined interest group issues concerning regulation of public water supplies.

eventually prospered by developing strong informal norms of trust, honesty, sobriety, and cooperation within their own community. These norms lowered their internal transaction costs substantially, permitting them to trade over space and through time in ways that outsiders in the broader Russian society could not. By the mid nineteenth century, in relatively unregulated sectors like the textile industry, they were the leading entrepreneurs in Russia. When the government's direct regulatory role increased in the late nineteenth century, their role in this sector then diminished (Danila Raskov 2002).

In the Russian economic system, informal norms evolving over long periods of time had fundamental consequences for who did what, and how efficiently. Excluded from formal state contract enforcement, the Old Believers developed their own informal system of transacting and enforcement. Later state regulations that elevated the importance of political connections, such as state licensing for entry, put them at a disadvantage.

These three historical examples illustrate the broad range of responses to regulation that arise from interest groups rather than from individual action. New regulations create new interests and destroy old interests. The rise and demise of shared interests can change the original politics irreversibly. The overall response to a regulation depends only partially on responses at the level of resource allocation. How the various interest groups respond to the regulation, and which ones seize the regulatory mechanisms over time—these are very significant issues. At this stage in our knowledge, these outcomes are extremely difficult to predict.

5. THE PERSPECTIVE OF NEW INSTITUTIONAL ECONOMICS

The standard neoclassical model in economics assumes that all transaction costs are zero. As Ronald Coase has established (1960), in such a world, if a non-equilibrium money price were set by regulation, the parties to an exchange could simply contract around the regulation. They could use vertical integration, other media of exchange, barter, long-term contracts, or alteration in product attributes. Even with money prices assigned randomly or set to zero, the system could work efficiently. The parties could price-discriminate perfectly, and no mutually advantageous trades would go unexercised (Barzel 1997). In that world, monopolies would not lead to deadweight losses, and there would be no externalities.

Stigler (1971) has commented that the zero-transaction-cost world is a strange world. Coase views the world of zero transaction costs as a useful first step in analysis, but he encourages us to limit the time we spend studying that world. He encourages us instead to focus more on the world of positive transaction costs, the world in which we live.

The standard literature on regulation typically modifies the neoclassical model slightly by assuming that positive (sometimes infinite) transaction costs exist along one or two specified margins, for example, “imperfect capital”

or “imperfect information,” while along all other margins they remain zero. These regulatory models pay lip service to utility maximization and opportunity costs—as opposed to purely pecuniary maximization and money prices—but usually ignore them in empirical work.

New institutional economics modifies the standard neoclassical model significantly by taking the view that transaction costs are positive and finite along all margins. This perspective draws on the fundamental tenet of economics: the relevant prices facing an individual are the full opportunity costs associated with the various choices. For a given good, these full opportunity costs include the money price of the good itself plus the transaction costs of obtaining the good.¹⁷

This perspective also emphasizes the possibility of substitution along many margins. Markets typically adjust through a mixture of changes in money prices and changes in other characteristics of the goods traded in that market, including location, waiting time, durability, warranty, freshness, and associated services. Indeed, in many markets, waiting time and quality are highly variable while money prices are relatively sticky.¹⁸ Inside households, firms, and even some markets, money prices may not be used at all.

This has many implications for measuring the full range of responses to regulation. The metric typically used—the money price—is an incomplete measure of the full range of responses possible. Attributes of the goods and also attributes of associated property rights, contractual forms, organizational structures, the medium of exchange, and informal networks—all of these can vary in response to new regulations. These responses can reduce or increase the deadweight losses traditionally associated with formal regulations.

The passage of time introduces additional considerations. A single snapshot of the regulatory process is not sufficient. The long-term impact of a regulation depends upon historical constraints and the nature of the evolving political and bureaucratic responses, including competition among interest groups. Over time, responses to regulations may alter the formation of political interest groups, political processes, and political outcomes. A regulation’s greatest impact may lie in its downstream consequences for political interest groups and for subsequent regulations. Inefficiencies created through these political and bureaucratic processes over time can easily be larger than the welfare losses measured by standard calculations.

As Coase (2002) and Krueger (1990) urge, we need many more careful case studies that are comparative in nature. These will eventually permit us to examine the multiple constraints facing decision-makers, to sort out general causal factors from historically specific factors, and to understand better the historical evolution of regulation and responses to it.

¹⁷The opportunity costs an individual faces are likely to be influenced by the specific characteristics of the individual, the type of exchange, and the institutional setting within which the individual is operating. See Benham and Benham (2001).

¹⁸Facing short- and long-term variations in market conditions, many restaurants deliberately alter waiting time and quality of service as well as money price.

A major objective in studying regulation is to understand its impact, the range of likely responses, and how these play out over time. The perspective of new institutional economics recognizes the role of history, path dependence, and the variability of experience across individuals, firms, and countries. A long-run task is to articulate a general approach that covers both economic and political actors, and both individual and collective action. A prudent scholar would examine policies with a keen eye on the process of creation and destruction of interests.¹⁹ If we can understand the potential consequences of regulators' actions in this broader sense, supplemented by detailed studies of specific regulations in many countries, we may do a better job of understanding the impact of regulation, responses to it, and the regulatory process itself.

ACKNOWLEDGEMENT

I thank Alexandra Benham for many discussions of these ideas and for her generous work in editing and re-editing this manuscript. Kurt Annen, Krishna Ladha, Marc Law, John Nye, Mary M. Shirley, Gili Yen, and two anonymous referees have provided helpful comments. Any errors remaining are mine.

REFERENCES

- Alchian, Armen and Reuben Kessel. 1962. "Competition, Monopoly, and the Pursuit of Pecuniary Gain" in *Aspects of Labor Economics*. National Bureau of Economic Research, Princeton, NJ: Princeton University Press, pp. 157–175.
- Alston, Lee, Gary Libecap, and Bernardo Mueller. 1999. *Titles, Conflict, and Land Use: The Development of Property Rights and Land Reform on the Brazilian Amazon Frontier*. Ann Arbor, MI: University of Michigan Press.
- Alston, Lee, Gary Libecap, and Bernardo Mueller. 2000. "Land Reform Policies, the Sources of Violent Conflict, and the Implications for Deforestation in the Brazilian Amazon". *Journal of Environmental Economics and Management* 39(2): 162–188.
- Bardhan, Pranab. 1997. "Corruption and Development: A Review of Issues". *Journal of Economic Literature* 35(3): 1320–1346.
- Barzel, Yoram. 1997. *Economic Analysis of Property Rights*, second edition. Cambridge: Cambridge University Press.
- Ben-Porath, Yoram. 1980. "The F-Connection: Families, Friends, and Firms and the Organization of Exchange". *Population Development Review* 6(1): 1–30.
- Benham, Lee. 1972. "The Effect of Advertising on the Price of Eyeglasses". *Journal of Law and Economics* 15(2): 337–352.
- Benham, Alexandra and Lee Benham. 1975. "Regulating Through the Professions: A Perspective on Information Control". *Journal of Law and Economics* 18(2): 421–448.
- Benham, Alexandra and Lee Benham. 2000. "Measuring the Costs of Exchange" in Claude Ménard (ed.), *Institutions, Contracts and Organizations: Perspectives from New Institutional Economics*. Cheltenham, UK: Edward Elgar, pp. 367–375.

¹⁹ Krishna Ladha (2002) has provided edification on this topic.

- Benham, Alexandra and Lee Benham. 2001. "Costs of Exchange". Accessed October 16, 2003 at <http://coase.org/w-benham2001thecostsofexchange.pdf>.
- Carlin, Wendy. 1989. "Economic Reconstruction in Western Germany, 1945: The Displacement of 'Vegetative Control'" in Ian D. Turner (ed.), *Reconstruction in Post-War Germany, British Occupation Policy and the Western Zones, 1945–55*, Oxford: Berg, pp. 37–65.
- Cheung, Steven N.S. 1975. "Roofs or Stars: The Stated Intentions and Actual Effects of a Rents Ordinance". *Economic Inquiry* 13(1): 1–21. Also in Lee J. Alston, Thrainn Eggertsson, and Douglass C. North (eds.), *Empirical Studies in Institutional Change*, Cambridge: Cambridge University Press, pp. 224–243.
- Cheung, Steven N.S. 1979. "Rent Control and Housing Reconstruction: The Postwar Experience of Prewar Premises in Hong Kong". *Journal of Law and Economics* 22(1): 27–54.
- Coase, Ronald. 1947. "The Origin of the Monopoly of Broadcasting in Great Britain". *Economica* 14: 189–210.
- Coase, Ronald. 1959. "The Federal Communications Commission". *Journal of Law and Economics* 2: 1–40. Also in the *Yearbook of Broadcasting Articles*, anthology edition. 1980. Washington DC: Federal Publication.
- Coase, Ronald. 1960. "The Problem of Social Cost". *Journal of Law and Economics* 3: 1–44.
- Coase, Ronald. 2002. Personal communication. October 23, 2002.
- de Soto, Hernando. 1989. *The Other Path*. New York: Harper & Row.
- Djankov, Simeon, Rafael La Porta, Florencio Lopez-De-Silanes, and Andrei Shleifer. 2002. "The Regulation of Entry". *The Quarterly Journal of Economics* 117: 1–37.
- Djankov, Simeon. 2003. "The Informal Economy: Large and Growing in Most Developing Countries". World Bank Hot Topic Discussion Board. Accessed October 16, 2003 at http://rru.worldbank.org/hottopics/Hot_Topics_djankov_2.asp.
- Gambetta, Diego. 1993. *The Sicilian Mafia: The Business of Private Protection*. Cambridge, MA: Harvard University Press.
- Kessel, Reuben. 1970. "The AMA and the Supply of Physicians". *Law and Contemporary Problems* 35(2): 267–283.
- Krueger, Anne O. 1990. "The Political Economy of Controls: American Sugar" in Maurice Scott and Deepak Lal (eds.), *Public Policy and Development: Essays in Honour of Ian Little*. Reprinted in Lee J. Alston, Thrainn Eggertsson, and Douglass C. North (eds.), *Empirical Studies in Institutional Change*, 1996. Cambridge: Cambridge University Press, pp. 169–218.
- Keynes, John Maynard. 1930. *A Treatise on Money, Volume 2: The Applied Theory of Money*. London: Macmillan.
- Ladha, Krishna. 2002. Personal communication, August 2002.
- Law, Marc. 2003a. *Specialization, Information, and Regulation in American Economic History*. Ph.D. dissertation, Washington University, St. Louis.
- Law, Marc. 2003b. "The Origins of State Pure Food Regulation". *Journal of Economic History* 63: 1103–1130.
- Lewis, C.A. 1924. *Broadcasting from Within*. London: George Nennes Limited.
- Marshall, Alfred. 1920. *Industry and Trade*, 3rd edn. London: Macmillan.
- Marvasti, Akbar and David Smyth. 1998. "Barter in the US economy: A Macroeconomic Analysis". *Applied Economics* 30: 1077–1088.
- Murphy, Kevin, Andrei Shleifer, and Robert Vishny. 1993. "Why Is Rent-Seeking So Costly to Growth?" *American Economic Review* 83(2): 409–414.
- Nee, Victor. 1998. "Norms and Networks in Economic and Organizational Performance". *The American Economic Review*, Papers and Proceedings 88(2): 85–89.
- Noll, Roger, and Bruce Owen. 1983. *The Political Economy of Deregulation: Interest Groups in the Regulatory Process*. American Enterprise Institute for Public Policy Research.
- North, Douglass. 1991. "Institutions". *Journal of Economic Perspectives* 5(1): 97–112.
- Olson, Mancur. 1982. *The Rise and Decline of Nations*. New Haven, CT: Yale University Press.

- Pollak, Robert A. 1985. "A Transaction Cost Approach to Families and Households". *Journal of Economic Literature* 23(2): 581–608.
- Radford, R.A. 1945. "The Economics of the P.O.W. Camp". *Economica* 12: 189–210.
- Raskov, Danila E. 2002. "Norms in the Economic Evolution: Old Believers in the Russian Nineteenth-century Textile Industry". Paper presented at the 2002 Conference of the International Society for New Institutional Economics, Cambridge, MA.
- Schneider, Friedrich. 2002. "Size and Measurement of the Informal Economy in 110 Countries Around the World". Paper presented at a Workshop of the Australian National Tax Centre, Canberra, Australia, July 17, 2002. Accessed October 16, 2003 at http://rru.worldbank.org/documents/informal_economy.pdf.
- Schranz, Mary. 1993. "Takeovers Improve Firm Performance: Evidence from the Banking Industry". *Journal of Political Economy* 101(2): 299–326.
- Shirley, Mary M. (ed.). 2002. *Thirsting for Efficiency: Experiences in Reforming Urban Water Systems*. London: Elsevier.
- Shleifer, Andrei and Robert Vishny. 1993. "Corruption". *Quarterly Journal of Economics* 108(3): 599–617.
- Singerman, Diane. 1995. *Avenues of Participation: Family, Politics, and Networks in Urban Quarters of Cairo*. Princeton.
- Smith, Adam. 1776. *The Wealth of Nations*. Chicago, IL: University of Chicago Press, 1976 edition.
- Stigler, George. 1951. "The Division of Labor Is Limited by the Extent of the Market". *The Journal of Political Economy* 59(3): 185–193.
- Stockder, Archibald H. 1932. *Regulating an Industry: The Rhenish-Westphalian Coal Syndicate 1893–1929*. New York: Columbia University Press.
- Sykuta, Michael. 1994. "Real Effects of Futures Markets on Firm and Industry Behavior: A Study of Institutions and Contracting in the Crude Oil Industry". Ph.D. dissertation, Washington University, St. Louis.
- Sykuta, Michael. 1996. "Futures Trading and Supply Contracting in the Oil Refining Industry". *Journal of Corporate Finance: Contracting, Governance and Organization* 2(4): 317–334.
- Wallis, John J. and Douglass C. North. 1986. "Measuring the Transaction Sector in the American Economy" in S.L. Engerman and R.E. Gallman (eds.), *Long Term Factors in American Economic Growth*, Chicago, IL: University of Chicago Press.
- Williamson, Oliver. 1985. *The Institutions of Capitalism*. New York: The Free Press.
- Williamson, Oliver and Scott Masten (eds.). 1995. *The Economics of Transaction Costs*. Northampton, MA: Edward Elgar.
- Williamson, Oliver. 1999. "The New Institutional Economics: Taking Stock/Looking Ahead". *Newsletter of the International Society for New Institutional Economics* 2:2.
- Yen, Gili and Lee Benham. 1986. "The Best of All Monopoly Profits is a Quiet Life". *Journal of Health Economics* 5(4): 347–353.