

NBER WORKING PAPER SERIES

THE RISE OF THE REGULATORY STATE

Edward L. Glaeser
Andrei Shleifer

Working Paper 8650
<http://www.nber.org/papers/w8650>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2001

We are grateful to Simeon Djankov, Claudia Goldin, Oliver Hart, Simon Johnson, Larry Katz, Louis Kaplow, Rafael La Porta, Florencio Lopez-de-Silanes, Steven Shavell, and Lawrence Summers for helpful comments, and to the NSF for funding. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2001 by Edward L. Glaeser and Andrei Shleifer. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Rise of the Regulatory State
Edward L. Glaeser and Andrei Shleifer
NBER Working Paper No. 8650
December 2001
JEL No. K2, K13, K42, L51, N41

ABSTRACT

During the Progressive Era at the beginning of the 20 th century, the United States replaced litigation by regulation as the principal mechanism of social control of business. To explain why this happened, we present a model of choice of law enforcement strategy between litigation and regulation based on the idea that justice can be subverted with sufficient expenditure of resources. The model suggests that courts are more vulnerable to subversion than regulators, especially in an environment of significant inequality of wealth and political power. The switch to regulation can then be seen as an efficient response to the subversion of justice by robber barons during the Gilded Age. The model makes sense of the progressive reform agenda, and of the successes and failures of alternative law enforcement strategies in different countries.

Edward L. Glaeser
Harvard University
315A Littauer Center
Cambridge, MA 02138
and NBER
eglaeser@harvard.edu

Andrei Shleifer
Harvard University
M9 Littauer Center
Cambridge, MA 02138
and NBER
ashleifer@harvard.edu

1. Introduction.

Before 1900, significant commercial disputes in the United States were generally resolved in courts. Courts ruled on liability in industrial accidents, on anti-competitive practices such as railroad rebates, on safety of foods and medicines, and even on the constitutionality of the income tax. In the three decades between 1887, when Congress passed the Interstate Commerce Act, and 1917, when “participation in the war put an end to the Progressive movement,” (Hofstadter, 1955), this situation changed radically. Over thirty years, reformers eroded the 19th century belief that litigation was the only appropriate response to the socially harmful acts. During the Progressive Era, regulatory agencies at both the state and the federal level began to replace courts in anti-trust policy, railroad pricing, food and drug safety, and many other areas. At the same time, the U.S. politics experienced other important changes, such as reform of the civil service, use of voter referendums to decide local issues, direct election of senators, recall of judges, and the growth of government more generally. In this paper, we attempt to understand why these changes occurred in the United States between 1887 and 1917.

Conventional economic theories of regulation do not help understand the Progressive movement. The standard public interest theory holds that regulation deals with market failures and externalities (Pigou 1934, Stiglitz 1989), but does not explain why either contract or tort law could not address these problems in the first place. Coase (1960) draws attention to this failure of regulatory economics to deal with the alternative possibility of judicial enforcement of laws and contracts, but does not address the choice of courts versus regulators. Posner (1992) and Shavell (1984a) discuss this choice between courts and regulators from the efficiency perspective. Posner (1992) emphasizes the fixed cost of lawsuits as a potential argument for regulation, whereas Shavell (1984a) points to the limits on the violator’s ability to pay as a drawback of litigation. Both of these theories predict that as cases became larger and the defendants’ pockets deeper

during the Gilded Age, efficiency calls for more litigation and less regulation. The reality, of course, was the opposite. Finally, Stigler (1971) and McChesney (1987) rejuvenated the economic theories of regulation by questioning the motives and the capabilities of the regulators. Their theories, however, are not helpful in thinking about the Progressive era, when the pressure for regulation came from the public and populist politicians rather than industry itself.

To discuss progressive reforms, we present a model of choice of a law enforcement regime between court-enforced liability and regulator-enforced input controls. A society can impose liability payments for accidents, for example, or require precautionary investments. Whatever approach a society chooses, law enforcement is not free, and the incentives, costs and opportunities facing law enforcers crucially shape the efficiency of alternative choices. As a consequence, different institutional arrangements are appropriate for different countries at different stages of economic and political development.

In our model, the crucial difference between liability and regulation as alternative systems of controlling market behavior is their vulnerability to subversion by the potential violator. Subversion includes such techniques as intimidating judges and regulators, bribing them, and using delay tactics to postpone a trial or a liability payment. By expending sufficient resources on subversion of justice, the potential violator can avoid either regulatory compliance or a liability payment. However, because liability regimes entail large payments with a small probability, such regimes are more vulnerable to ex post subversion than the regulation of inputs, especially in countries lacking law and order in the first place. Moreover, the problem of subversion of liability regimes becomes worse in situations of extreme inequality of economic and political resources between plaintiffs and defendants. In such circumstances, regulation is more efficient than liability. This reason for the switch to regulation complements several others

of relevance to progressive reform, such as the stronger incentives and greater specialization of regulators compared to judges (Landis 1938, Glaeser, Johnson, and Shleifer 2001).

The theoretical analysis points to a fundamental change that made it efficient for the American society to increasingly rely on regulation. The commercialization and industrialization of the economy in the second half of the 19th century created extreme inequality of resources among the parties in tort and contract disputes, and therefore gravely exacerbated the problem of subversion. When courts and legislatures could be bought or manipulated, individuals and small companies were unlikely to prevail against “robber barons” in accident, restraint of trade, or discriminatory tactics disputes. From our perspective, the regulation of markets was first and foremost a response to the problem of subversion of law enforcement through courts. Moreover, the other political changes of this period, such as the Civil Service reform, increasing importance of direct elections, and judicial recalls can also be understood from the perspective of controlling subversion.

Woodrow Wilson (1913) held the failure of courts to be a crucial reason for the necessity of regulation in his *New Freedom* program.

The judicial process is the ultimate safeguard of the things that we must hold stable in this country. But suppose that that safeguard is corrupted; suppose that it does not guard my interests and yours, but guards merely the interests of a very small group of individuals; and whenever your interest clashes with theirs, yours will have to give way, though you represent ninety percent of the citizens, and they only ten percent. Then where is your safeguard? (Wilson, 1913, p. 240).

Writing during the New Deal, Landis (1938) saw regulation as a political response to the failure of courts to keep up with the community ideas of justice. He thought that the advocacy of “leaving the problems of railroad charges and management to work themselves out in the courts as questions arise from time to time...indicates a singular unawareness of the fact that the chief drive for the resort to the administrative process in the field of railroad regulation arose from a

recognition that the remedies that the courts could provide were insufficient to make effective the policies that were being demanded.” In effect, our model provides an efficiency rationale for this switch from litigation to regulation.

In addition to helping clarify a crucial period of U.S. economic theory, the model sheds light on a number of theoretical and policy problems of law enforcement. Economists since Coase (1960) have been interested in the question of whether regulation or litigation is a better way to deal with tort problems. Coase argued generally that “transaction costs” should determine the answer, and the successor literature has identified a range of such costs (e.g., Shavell 1984a, 1984b). Here we point to a different kind of transaction costs, namely the ability of the violators to subvert a law enforcement mechanism that imposes on them a very high penalty, and argue that this force is crucial for the choice of law enforcement strategy. This approach also helps us understand why Becker’s (1968) “boil them in oil” enforcement strategies, which entail very high penalties with a very small probability and hence minimize investigation costs, are rarely workable: they will be subverted *ex post*.²

More generally, the analysis sheds light on a critical determinant of optimal law enforcement strategy, namely how vulnerable the system of justice is to subversion. In situations of extreme vulnerability to corruption or intimidation, the optimal strategy might be no legal or regulatory restrictions at all, as the alternative is the socially costly regime in which rules are simply subverted. In the intermediate regimes, the most efficient policy is regulation. Although it is not as efficient as litigation *absent subversion*, regulation is less vulnerable to subversion

² This reasoning suggested to us an alternative title for the paper, “The Denise Rich effect” in recognition of a friend of President Clinton who secured a pardon for her former husband by making a donation of a few million dollars to the Clinton Presidential library. This event apparently followed an offer to the U.S. government from Mr. Rich, a fugitive financier, of \$100 million to settle all charges against him, which was turned down by the U.S. Justice Department that wanted him to stand trial and go to jail. The point here, as in our model, is that harsh penalties lead to subversion of justice.

than litigation, and might be socially preferred for that reason. Finally, in the regimes where the system of justice is least vulnerable to subversion, liability regimes – either strict liability or negligence – become the optimal strategy. In broad terms, the paper suggests that the optimal strategy of law enforcement depends on how much order the country has in the first place.

In the next section, we sketch the U.S. regulatory situation in the second half of the 19th century, and the changes it underwent during the Progressive era. In Section 3, we present a model of the choice of a law enforcement strategy. In Section 4, we apply the model to the U.S. experience, while in Section 5 we draw broader implications. Section 6 concludes the paper.

2. Enforcement by courts in the 19th century

We start with three claims about the United States during the “Gilded Age” between the Civil War and the Progressive Era. First, until the end of the 19th century, the U.S. in many respects followed the laissez-faire ideal in which courts rather than regulatory agencies dealt with socially harmful acts. Second, after the Civil War, wealth and power regularly subverted the workings of the court system. Third, this subversion of courts entailed outcomes radically different from those suggested by Coase’s (1960) benign vision of judicial enforcement. Traditional arguments for the failure of the Coase theorem, such as transaction costs (Shavell 1987), do not explain the limits of the late 19th century American justice. Instead, 19th century courts failed because money and power regularly influenced the path of justice.

By the late 19th century, tort law became one of the great edifices of the common law, whose development was greatly accelerated by the industrial revolution in general and railroads in particular. “Trains were also wild beasts; they roared through the countryside, killing livestock, setting fires on houses and crops, smashing wagons at grade crossings, mangling passengers and freight. Boilers exploded; trains hurtled off tracks; bridges collapsed;

locomotives collided in a grinding scream of steel. Railroad law and tort law grew up, then, together. In a sense, the two were the same (Friedman 1985, p. 468).” In cases of both personal and social harm, individual sought damages in common law courts. In theory at least, legal doctrine of the 19th century was more strongly oriented towards eliminating externalities than modern doctrine. Horwitz (1992) describes how the maxim *sic utere*, “use your own so as not to injure others,” was invoked by many common law judges. This maxim justified court action against a variety of perceived nuisances (saloons, gunpowder storage facilities, slaughterhouses) “which could be abated without any justification of the defendant.”

In the development of tort law, the 19th century saw the great debate between strict liability and negligence standards (Horwitz 1992).³ The English case of *Rylands v. Fletcher* (1868), in which a court held a landowner liable when a reservoir he built on his own property accidentally flooded the plaintiff’s coal mine, stands as the great statement of strict liability. In contrast, American jurists, most notably Holmes (1881), sought to make negligence a more common rule. Several great legal reputations of this period, including those of Holmes, Cardozo and Brandeis, are intimately linked to their contributions to the development of tort law.⁴

While the courts developed the law, they also protected their dominance in this area. Supreme courts regularly struck down nascent attempts at regulation by appeals to freedom of contract. In *Lochner v. New York* (1905), the New York and U.S. supreme courts ruled that maximum hours legislation was unconstitutional for bakeries because “there is no reasonable ground for interfering with the liberty of person or the right of free contract, by determining the

³ Strict liability means that the plaintiff just needs to show causation. Negligence means that the plaintiff needs to show both causation and some form of negligence.

⁴ Schwartz (1981, 1989) argues that the movement from strict liability to negligence is not detectable prior to the Civil War. Our theory focuses on the post-civil-war period, where the change is clearer.

hours of labor, in the occupation of a baker.” In the Gilded Age, courts provided remedies for harm. Regulation was rare and regulatory agencies almost non-existent.

Despite the dominance of the courts, many observers of the Gilded Age see them as ineffective and partial arbiters of justice. Skocpol (1992) complains about the pro-business attitudes of the courts. Horwitz (1992) argues that this era is replete with pro-business jurists, and associates the negligence principle with attempts by pro-business jurists to protect firms from lawsuits. Lockard and Murphy (1992) claim that judges supported corporations because of “a campaign to ‘educate’ judges about the sacredness of private property.” According to Friedman (1985), “What they [the leading concepts of tort law] added up to was also crystal-clear. Enterprise was favored over workers, slightly less so over passengers and members of the public.... The thrust of the rules, taken as a whole, approached the position that corporate enterprise should be flatly immune from actions for personal injury (p. 475).”

Judicial ideology, of course, was itself shaped by the selection of judges, and through their influence on the political machines, industrialists were able to determine who sat on the bench. According to Woodrow Wilson (1913, p.242), “The disease lies in the region where these men [judges] get their nominations; and if you can recover for the people the *selecting* of judges, you will not have to trouble about their recall. Selection is of more radical consequence than election.” Callow (1966, p. 135) writes that “the appointment of the right men to key posts is the third step in making the law as much a matter of politics as of justice.”

After the selection, judges as well as prosecutors were influenced by the rich and powerful through both legal and illegal means. Nineteenth century corporations projected substantial political influence, superior lawyers and ready access to large legal war chests. Their lawyers produced briefs that exonerated their clients and slowed down the wheels of justice for years. Last but not least, they routinely bribed legislators, judges, and juries.

Lloyd (1894) describes how Standard Oil subverted the attempts of its opponents to secure damages in courts by intimidation and bribery of witnesses, payments and political pressures on judges and legislators, and theft and destruction of evidence. Tarbell (1903) corroborates the outline of Lloyd's account. In "Robber Barons," Josephson (1934) relates the story of the battle for the Erie railroad between Commodore Vanderbilt and Jay Gould. The battling barons acquired a number of judges and legislators, who issued laws, rules and injunctions preventing each party from exercising their powers over the railroad on request. The battle culminated in an open auction of policies by the New York legislature, which Gould won by paying higher bribes to more legislators.

The subversion of justice is a pervasive theme of the muckrakers. In "Tweed Days in St. Louis," Lincoln Steffens tells the story of a young Circuit Attorney, Mr. Folk, who was put on the Democratic ticket by mistake, and upon election began putting both Democratic and Republican officials behind bars for corruption. Having failed to persuade Mr. Folk to cease and desist, the local political machine decided to get rid of him. "At the meeting of corruptionists three courses were decided upon. Political leaders were to work on the Circuit Attorney by promise of future reward, or by threats. Detectives were to ferret out of the young lawyer's past anything that could be used against him. Witnesses would be sent out of town and provided with money to remain away until the adjournment of the grand jury." In Minneapolis, Steffens describes the coercion of the jury in the trial of the corrupt Mayor Ames. He points to the foreman of the jury, "to whom \$28,000 was offered to quit, and for whose slaughter a slugger was hired to come from Chicago" (Steffens, 1906).

Albert Cardozo—the father of the great progressive jurist Benjamin Cardozo—is a classic example of a judge in league with a political machine. Appointed by William Marcy Tweed to the New York State Supreme Court, Albert Cardozo served Tweed ably. "He became

a kind of escape hatch for criminals; through his good offices he pardoned or dismissed several hundred known criminals who were or might be useful to the [Tweed] Ring (Callow 1966).” At his own impeachment hearings, Cardozo was accused of crimes ranging from standard nepotism (his nephew was a beneficiary of his power) to releasing convicted clients, to helping Fisk and Gould reduce the losses from their Gold Conspiracy through the legal system.

The pervasive distortion of justice through legal and illegal forms of influence decided many cases and had a broad influence on the 19th century economy. Courts often failed to address the grievances of the parties damaged in the new economy, such as workers suffering from accidents, producers suffering from abusive tactics by the railroads, or consumers poisoned by bad food, and ruled in favor of large corporations. Tort claims against railroads, following large scale accidents that often killed third parties, proved to be slow and often unsuccessful. In *Ryan v. New York Central Railroad* (1866), the New York court of appeals argued that even if a railroad caused a fire, its liability extended only to the immediately adjacent house and not to other homes destroyed by the fire. The court held that “to sustain such a claim...would subject [the railroad] to a liability against which no prudence could guard, and to meet which no private fortune could be adequate” (Friedman 1985, p. 469). More generally, the courts’ view of such accidents was “NOBODY TO BLAME” (Friedman 1985, p. 470).

Fishback and Kantor (2000) describe the extreme inequality between firm and employee in industrial accident claims. The legal strength of corporations meant that many injured families settled for relatively small amounts of money. Summarizing a number of studies, Fishback and Kantor estimate that families of workers killed in industrial accidents received an average of 8 months pay, and nothing in about 40% of the cases (p.34). Under political (and perhaps more direct) pressure from manufacturers, courts adopted the position that the estate of

an injured worker must prove negligence by the company, and that evidence of negligence by the worker himself, or by one of his co-workers, absolved the company of liability.

In sum, the system of dispute resolution by courts, which emerged in agrarian America in the 18th and early 19th centuries, was not suited for the conditions of the late 19th century. The reason for this failure is that large corporations possessed economic resources far in excess of those at the disposal of their opponents – whether individuals or small firms – and could use these resources to subvert justice. The mechanisms of subversion ranged from superior legal talent to political pressure to outright bribery. For our purposes, the exact mechanism does not matter. What matters is that courts did not make the perpetrators pay for the social harm of their actions. As a consequence, the system broke down.

The Regulatory Response

As the courts supported the corporations, political actors sought a different response to social harm. Starting with Charles Francis Adams and the Massachusetts Railroad Commission, regulatory agencies became a substitute for judicial action. The reforms started with states and municipalities, but eventually moved to the federal level. In 1887, the Interstate Commerce Act created the Interstate Commerce Commission, which – combined with subsequent legislation – had the power to stop railroad rebates and ultimately, under the Hepburn Act of 1906, to set rates. In 1890, Congress passed the Sherman Act, restricting the formation of trusts.

The real growth of regulatory activity occurred during the Progressive era after the turn of the century, under the presidencies of Theodore Roosevelt and Woodrow Wilson. In 1906, under muckraking pressures, Congress passed Pure Foods and Drugs to control the distribution of medicines, as well as a federal meat-inspection law. Under Wilson, banking came under federal regulation in the Federal Reserve Act of 1913 and anti-trust enforcement was intensified

under the Clayton Act in 1914, to name just two measures. Fishback and Kantor (2000) take the adoption of workers compensation laws by several U.S. states around the turn of the century to be a response to the failure of courts to address the problems of workers injured or killed in industrial accidents.

Although the growth of regulation slowed down, and arguably retreated, under the Republican presidencies of the 1920s, it revived and accelerated under Franklin Roosevelt in the 1930s. Among the most notable measures of this period were the Securities Acts of 1933 and 1934, written in part by Landis, which brought securities markets under federal regulation. When Landis wrote in 1938, he could confidently conclude that “the administrative” has replaced “the judiciary” as the principal form of social control of business.

3. The Model.

A firm can take a level of precaution, equal to Q_1 or Q_2 , in order to avoid an accident. The high level of precaution, Q_2 , requires a cost of S times C , while the low level of precaution, Q_1 , is free. The parameter C represents the cost of precaution per unit of production—the level of S represents the scale of the firm. There are two types of firms that might be responsible for the accident: α 's and β 's. For α types the probability of an accident is unaffected by precaution and equals P_α . This implies, importantly, that it is not efficient for type α firms to invest in a high level of precaution. Let π_α be a proportion of firms that are of type α . For type β firms the probability of an accident equals P_1 or P_2 , with $P_1 > P_2$, depending on whether the level of precaution is Q_1 or Q_2 . The accident imposes a social cost of S times D , where D refers to the social cost per unit of economic activity and S refers again to the scale of the firm.

For our purposes, it is not important whether D is concentrated (as in the case of a workplace accident) or widely shared (as in the case of pollution). When the damage is concentrated, simple ex ante contracts can sometimes deal with possible damages, even without a liability regime. For example, an employee working in a dangerous occupation would receive a higher wage, which compensates him for the risk of an accident. Even in such situations, however, it is generally efficient for the firm to provide the risk-averse employee with some accident insurance (Shavell 1987). As a consequence, the firm has an incentive to avoid paying after the accident takes place. To the extent that the dispute needs to be settled in court, as such issues often are (Fishback and Kantor 2000), our analysis remains pertinent. We make:

Assumption 1: $(P_1 - P_2)D > C$,

so that for the β types, unlike for the α types, the high level of precaution is socially valuable.

We consider the standard menu of possible regulatory schemes: (1) strict liability, (2) negligence, and (3) regulation of inputs.

1. Strict liability requires that, in case of an accident, the firm must pay a fine any time that damages occur.
2. Negligence means that a fine is charged whenever damages occur and when the firm undertakes the low level of precaution.
3. Regulation requires that the high level of precaution be taken and imposes a fine, F , whenever the firm fails to do so. We assume that the regulator cannot distinguish between the two types of firms.

With no problems of enforcement, strict liability can achieve the first best as long as the fine (denoted by F) is greater than $\frac{SC}{P_1 - P_2}$. Negligence can also achieve the first best whenever

$\frac{SC}{P_1} < F < \frac{SC}{P_\alpha}$. Note that $\frac{SC}{P_1} < \frac{SC}{P_1 - P_2}$. Because strict liability uses less information, it

requires a larger fine. At the same time, strict liability provides no incentives for the type α firms to invest in the useless for them high level of precaution, and therefore eliminates this source of inefficiency.

Finally, we assume that the regulator learns that the firm is using low levels of precaution with exogenous probability p . We have also considered the case where p is endogenous. Since it yields qualitatively similar results, we focus here on the simpler scenario with exogenous p . To allow regulation to be a potentially efficient scheme in the presence of subversion we make:

Assumption 2: $p > P_1$.

The idea behind this assumption is that the regulation can be designed so that detection of the failure to invest in precaution is relatively inexpensive and certain. This could involve mandating well-specified “bright line” rules, such as the use of safety equipment or of warning labels for customers. In addition, some regulations are designed to encourage third party enforcement. For example, employees or customers themselves can occasionally cheaply identify violations and complain (Glaeser and Shleifer 2001a).

We turn next to optimal law enforcement in the presence of subversion. We assume that the firm can impose a cost of A on the law enforcer who wants to punish it. This cost may reflect character assassination, complaints to superiors, intimidation, delay, or political contributions to opponents. The cost to the firm of imposing A on the law enforcer is given by $K(A)$, which we take to be a social as well as a private cost. A downward shift of the function $K(A)$ can be interpreted as an improvement in the technology to subvert justice available to the firm, so we think of regimes with a higher $K(A)$ as possessing higher levels of “law and order.”

For simplicity, we assume that the firm must invest in $K(A)$ up front, before any accidents occur or decisions as to its liability are taken by law enforcers. This can be thought of as an investment in the capacity to impose pain (“armaments”). Once this capacity is acquired, we assume that the marginal cost of imposing pain A on the law enforcer is zero, so that the firm is always willing to do so when it does not like the decision. This formulation avoids the credibility issues inherent in most crime and punishment models, where the violator might not want to follow through with punishing the law enforcer (and vice versa). This formulation also avoids the possibility of the firm bargaining with the law enforcer over the verdict and the bribe. However, this model better describes intimidation and legal tactics than bribes per se.

The benefit to a law enforcer such as a judge or a regulator of enforcing the law equals B , which we take to be exogenous. It can reflect the degree of judge/regulator’s career concerns or innate integrity (Posner 1995). In this paper, we do not consider the differences in incentives between judges and regulators, which might be important (Glaeser, Johnson, and Shleifer 2001). If $A \geq B$, the violator subverts justice and avoids punishment. In this case, the violator rationally chooses $A=B$ and incurs a cost of $X \equiv K(B)$. The possibility of subversion of justice means that if the fine is greater than X , the firm will choose to subvert justice and pay X instead.

Recall that we have assumed (Assumption 2) that $\frac{C}{p} < \frac{C}{P_1} < \frac{C}{P_1 - P_2}$, where $\frac{C}{p}$, $\frac{C}{P_1}$ and $\frac{C}{P_1 - P_2}$ are minimal fines that overcome the subversion of justice in regulation, contributory negligence and strict liability schemes respectively. We can now formulate

Proposition 1. If $P_\alpha < P_1$, then:

(a) for $\frac{X}{S} < \frac{C}{p}$ the only feasible option is laissez faire;

(b) for $\frac{C}{p} < \frac{X}{S} < \frac{C}{P_1}$ regulation dominates laissez faire if $D > \frac{C}{(1-\pi_\alpha)(P_1-P_2)}$, and vice versa if

this condition does not hold;

(c) for $\frac{C}{P_1} < \frac{X}{S} < \frac{C}{P_1-P_2}$ negligence achieves first-best; and

(d) for $\frac{X}{S} > \frac{C}{P_1-P_2}$ both negligence and strict liability achieve first-best.

If $P_\alpha > P_1$, then:

(a) for $\frac{X}{S} < \frac{C}{p}$ the only feasible option is laissez faire;

(b) for $\frac{C}{p} < \frac{X}{S} < \frac{C}{P_1-P_2}$ regulation dominates laissez faire if $D > \frac{C}{(1-\pi_\alpha)(P_1-P_2)}$ (negligence

achieves the same social outcomes but requires larger fines), and vice versa if this condition does not hold.

(c) for $\frac{X}{S} > \frac{C}{P_1-P_2}$ strict liability achieves the first-best.

For $P_\alpha < P_1$, Figure 1 illustrates the proposition graphically.

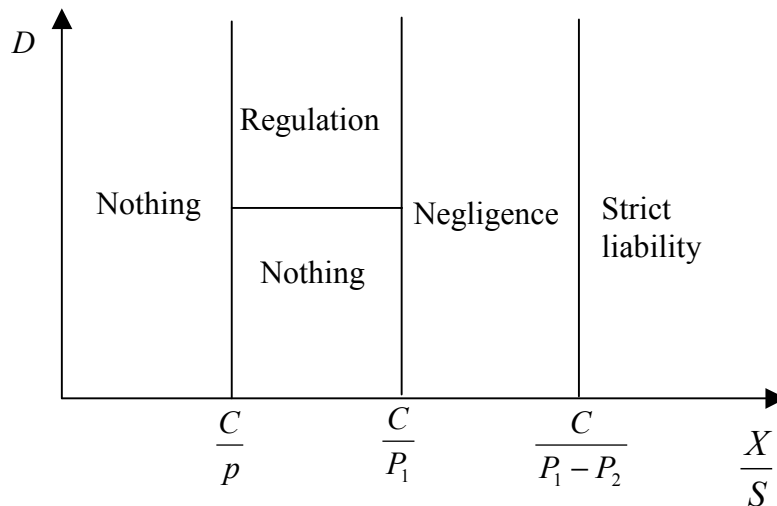


Figure 1

Proof. We consider four different cases with respect to the level of the cost of the subversion of justice, X/S , relative to the minimum fines necessary to implement alternative enforcement schemes (see Figure 2).

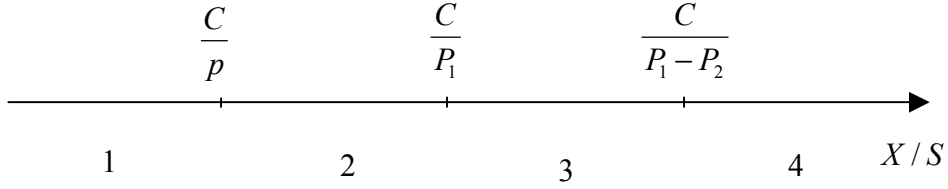


Figure 2

1. $\frac{X}{S} < \frac{C}{p}$. In this case, the firm has access to cheap intimidation technology, so no regulatory scheme can force the firm to invest in precaution. Any fine – from regulation, strict liability, or negligence, will be undone. The expected social loss if there is no regulation equals $L_N = \pi_\alpha P_\alpha SD + (1 - \pi_\alpha) P_1 SD$.
2. $\frac{C}{p} < \frac{X}{S} < \frac{C}{P_1}$. In this case, neither strict liability nor negligence schemes can implement efficient level of precaution for the type β firms. The two relevant options are to regulate or to do nothing. Regulation forces firms to maintain the high level of precaution if the fine, F , is set in the interval $\frac{SC}{p} \leq F < X$. Under our assumptions, both types of the firms choose the high level of precaution, which is excessive for type α firms. The expected social loss under regulation equals $L_R = \pi_\alpha P_\alpha SD + (1 - \pi_\alpha) P_2 SD + SC$. When C is small, i.e. $(1 - \pi_\alpha)(P_1 - P_2)D > C$, then excessive precaution is less wasteful than the expected damage from its absence, and regulation dominates laissez faire regime. When $C > (1 - \pi_\alpha)(P_1 - P_2)D$, then doing nothing is more efficient.

3. For X/S in the interval $\frac{C}{P_1} < \frac{X}{S} < \frac{C}{P_1 - P_2}$ negligence becomes a subversion-of-justice-

proof option. If $P_\alpha < P_1$ then negligence achieves the first best outcome when F is set in

the interval $\frac{SC}{P_1} < F < \min\{X, \frac{SC}{P_\alpha}\}$. The expected loss in this case is equal to

$$L_{CN} = \pi_\alpha P_\alpha SD + (1 - \pi_\alpha)(P_2 SD + SC) = L_{FB}.$$

If $P_\alpha > P_1$ then negligence cannot achieve both the efficiently low level of precaution for the type α firms and the efficiently high level of precaution for the type β firms. The social loss coincides with that under regulation, and the choice between laissez faire and regulation is the same as in case 2.

4. If $\frac{X}{S} > \frac{C}{P_1 - P_2}$, then the subversion of justice is not an acute problem, and high fines

necessary for strict liability scheme cannot be avoided. Any fine larger than $\frac{SC}{P_1 - P_2}$ and

smaller than X leads to the first best outcome in the strict liability scheme.

In sum, among the three possible regulatory schemes, only negligence and strict liability can in principle achieve the first best outcome. Strict liability, to be feasible, requires high cost of the subversion of justice, but has the advantage that it does not incentivize firms that do not benefit from precaution to invest in it. Negligence is less prone to subversion of justice, but does distort the incentives of the firms that do not benefit from it.

Perhaps the principal message of this model is the tight relationship between the “law and order” already prevailing in a society, and the optimality of alternative law enforcement schemes. As Figure 1 illustrates, in the environment of low law and order, doing nothing might

be superior to imposing legal and regulatory rules that are only going to be subverted at some social cost. In either case, harmful conduct is not punished, but with *laissez-faire*, corruption and other forms of subversion are avoided. In the regime of intermediate law and order, regulation might be efficient precisely because the fines it entails to achieve compliance are modest, and therefore will not be avoided through subversion of justice. The efficiency of regulation in this regime comes precisely from the fact that the penalties associated with the liability regime are too high, and therefore in such a regime justice is subverted. It is more efficient to have the cruder but enforceable system of regulation, than the finer and unenforceable system of liability. Finally, after high levels of law and order are already achieved, the society can attempt to resolve disputes through courts, using either a contributory negligence or a strict liability system. These systems call for the imposition of high fines *ex post*, and are therefore only feasible when the liable parties can be compelled to pay these fines without subverting justice.

The case for optimality of regulation suggested by this model is related to, but different from, those considered in our earlier work. Glaeser, Johnson, and Shleifer (2001) follow Landis (1938) and present a model in which it is easier for the society to provide high powered incentives for regulators than for judges. In the terminology of this model, B is higher for regulators than for judges. This too provides a rationale for regulation in some circumstances, especially when the costs of identifying violations are high. Glaeser and Shleifer (2001a) do not deal with the incentives of law enforcers, but rather with the costs of identifying violations. The paper argues that, in some circumstances, regulatory “bright line rules” make it cheaper to identify and penalize violations than extracting damages for harm or taxing undesirable conduct. For example, blue laws forbidding liquor sales on Sundays may be easier to enforce than Sunday liquor taxes, especially when private detection of violations is part of enforcement.

Finally, in Glaeser and Shleifer (2001b), we also consider the problem of subversion of justice, but focus on the historical development of common and civil law. We argue that the civil law institutions of state-employed professional judges following clear legal rules under constant superior review developed in France as a response to high levels of disorder prevailing in the society. These institutions were a mechanism of protecting law enforcers from subversion by powerful litigants. In contrast, the common law institutions of adjudication by lay juries following broad principles of community justice with only perfunctory appeal were more appropriate for the relatively more orderly England, where subversion of justice by the powerful magnates was a less pervasive problem. Although this paper and Glaeser and Shleifer (2001b) are clearly cousins, they deal with very different consequences of the problem of subversion.

4. Implications of the Model for the Progressive Era.

We can use the model of the previous section to understand the rise of regulation in the United States at the end of the 19th and the beginning of the 20th centuries. Our interpretation is that “S”—the scale of economic activity—rose dramatically over the 19th century. During the industrial revolution, firms grew sharply in size. The social costs of harm grew roughly proportionately, but the costs of subverting justice did not. As a result, a legal system that may have operated well during the agrarian period, failed when faced with entities that had huge incentives to subvert it. Because higher levels of “S” lead to subversion of both strict liability and negligence, regulation became the efficient response.

During the first half of the 19th century, the bulk of the economy was agricultural; 64 percent of workers in 1850 were in farm occupations. The manufacturing that did exist was concentrated in small firms. The McLane report, a large but incomplete survey of the economy, found only 106 manufacturing firms in 1832 with assets above \$100,000. Chandler (1977) saw

production in the mid-19th century as “being carried out by a large number of small units employing less than fifty workers...”

When a firm caused social harm in 1830, the judicial system could usually deal with it. Assessed damages would generally be small and the firm did not have a strong incentive to subvert justice. Moreover, the firm itself was small and did not have the resources to subvert the system. We take this to mean that the U.S. in 1830 is described by the right part of Figure 1, where $X > \frac{SC}{P_1 - P_2}$. In that region, strict liability produces the first best outcome.

This situation changed after the Civil War. The building of railroads, the growth of industrial enterprises and mines, and the creation of large financial firms introduced into the American economy disputes between parties of highly differentiated economic and political resources. America became far less agricultural; by 1900, only 37 percent of workers were in farm occupations. The manufacturing industry grew more than tenfold (measured by number of employees) over this time period. The new technologies often proved to be deadly. Thousands of passengers and third parties died in railroad accidents. By 1900, there were approximately 35,000 deaths and two million injuries annually in industrial accidents (Friedman, 1985).

New technologies also raised the scale of both railroads and manufacturing firms, as national markets developed. Innovations in organizational form made larger firm size feasible. Chandler (1977) identifies 278 firms with more than \$20 million in assets in 1917. Hofstadter (1955) writes that “the Morgan interests at the peak of the financial systems held 341 directorships in 112 corporations ... with aggregate resources or capitalization of \$22,245,000,000.” This vast expansion in firm size corresponds to an increase in the parameter “S” in the model.

The model corresponds most closely to situations where the industrial revolution raised both the precaution costs and the size of single accidents, e.g., train wrecks or factory fires (such as the famous Triangle Shirtwaist Fire, where 146 workers died). In other instances, such as industrial accidents involving individual workers, the cost of precaution and the number of accidents rise with enterprise size, but the size of each individual case remains small. If such cases are isolated, it might not pay a firm to subvert justice. In practice, however, the stakes in such cases are large for large firms, because of the power of legal precedents. When a steel mill or a railroad concedes liability in an industrial accident, it becomes vulnerable to claims in all the future similar accidents. As a consequence, even in such situations, the assumption of the model that the willingness to subvert justice rises with the scale of the firm is appropriate.

The model treats S as exogenous, but also points to an incentive for firms to increase it. In the regimes of either strict liability or negligence, rising values of S make no difference to corporate profits when corporations actually pay the fines, as long as fines scale with S . However, when X is independent of S , when firms subvert justice, higher values of S lead to lower payments for social harm. The payment per unit of economic activity is X/S and as such, larger firms pay less.⁵ Indeed, the advantage of scale in subverting the legal system is a competitive edge of 19th century corporation. The muckraking literature sees the largest firms as particularly effective in shaping legal outcomes.

The economic creations of the late 19th century, such as national railroad systems and trusts, may have been designed to gain political and not just economic power. The entrepreneurs

⁵ This is not so when disputes are between two entities possessing comparable abilities to subvert justice, as in the battle for the Erie railroad. In this case, a judge or a legislator can just auction off a decision, with the result that X actually scales up with S . It has been said that commercial law in Russia in the 1990s was developing through the disputes between the so-called oligarchs, where each side has the capacity to bribe a judge, and so must appeal to the law as well in order to win its case.

could use the economic and political resources that such combinations brought to buy both politics and justice. In a similar vein, and for similar reasons, financial-industrial groups were created in Russia in the 1990s with the basic goal of attaining political rather than economic influence (Nagel 1999).⁶

The progressives cited the rise in the scale of enterprise as a primary reason for new government action. In 1888, Charles Eliot noted that a modestly sized Boston railroad had three times as many employees as the Commonwealth of Massachusetts. The first chapter of Wilson's manifesto *The New Freedom* states that "the employer is now generally a corporation or a huge company of some kind," and as a result "new rules must be devised..." Wilson's progressive rival Theodore Roosevelt viewed Herbert Croly's *The Promise of American Life* as his intellectual sourcebook. Because of "the existing concentration of wealth and financial power in the hands of few irresponsible men," Croly believes "efficient regulation there must be." Progressives saw the scale of industrial enterprise as the root of the republic's problems.

The model helps us understand Croly's call for regulation. As S rises, the society moves leftward across Figure 1. First, negligence becomes optimal and should replace strict liability. We have already noted that, in the 1870s, negligence replaced strict liability as the standard in tort law. As S continues to rise, negligence itself fails to produce desirable results. When S is sufficiently high, and when D is also high, regulation provides the only reliable recourse against social harms. This is ultimately our explanation of why the progressives saw the need for

⁶ In the first volume of the *American Economic Review* in 1887, Henry Adams wrote as follows: "The policy of restricting public powers within the narrowest possible limits tends to render government weak and inefficient, and a weak government placed in the midst of a society controlled by the commercial spirit will quickly become a corrupt government; this in its turn reacts upon commercial society by encouraging private corporations to adopt bold measures for gaining control of government machinery. Thus the doctrine of laissez-faire overreaches itself; for the application of the rule which it lays down will surely destroy the harmony between public and private duties essential to the best results in either domain of action."

regulation. Because input regulation is more consistently applied and involves smaller penalties, it provides a weaker incentive for the subversion of justice. If government action is at all responsive to efficiency, the model predicts exactly the course of the Progressive era: increased government regulation of business should follow the increase in its scale.

The Progressives themselves saw the problems of the judiciary and advocated regulation. In *The New Freedom*, Wilson argues that it is “the right of government ... to see whether accidents were properly guarded against,” and that “if somebody puts a derrick improperly secured on top of a building or overtopping the street, then the government of the city has the right to see that that derrick is so secure that you and I can walk under it and not be afraid that the heavens are going to fall on us.” Why doesn’t Wilson just trust the tort system as administered by the judiciary? He also writes:

There have been courts in the United States which were controlled by the private interests. There have been supreme courts in our states before which plain men could not get justice. There have been corrupt judges; there have been controlled judges; there have been judges who acted as other men’s servants and not as servants of the public (Wilson, 1913, p. 240).

Croly also casts aspersions on judges and calls them “creatures of the political machine,” and advocates his own more ambitious plan of government ownership. As Hofstadter (1955) writes: “the average American tended more and more to rely on government regulation, to seek in governmental actions a counterpoise to the power of private business.”

Understanding the Progressive Program

The Progressive program was not merely a revolution in regulation. The law saw a switch from strict liability to negligence, which according to the model should follow an increase in scale. More generally, the model suggests that efficient reforms should aim to raise X or reduce S. Many progressive reforms can indeed be understood from this perspective.

One progressive innovation was the introduction of regulatory agencies to oversee specific areas of the economy and to punish socially harmful conduct.⁷ Glaeser, Johnson and Shleifer (2001) argue that the fundamental difference between courts and regulators lies in the incentives of regulators. In the common law tradition, judges are regularly protected from politicized incentives. Regulators, in contrast, are rewarded for finding violations. In the model, the incentives created by regulation increase B. Many progressives also supported the “recall” of government officials, especially judges. This innovation made it possible for voters to oust officials failing to perform as the public wanted. We see this as an attempt to raise B for judges, or alternatively to increase the cost they face if they give in to coercion, i.e. X.

Equally important to the progressives was the professionalization of bureaucracies, or civil service reform. Croly advocates a more widespread use of life tenure for judges, which would have a similar effect of reducing the political influence on them. The progressives’ simultaneous support of increased democracy in the form of judicial recall and of reduced democracy in the form of civil service reform is puzzling. Both reforms, however, can be understood as attempts to increase X. While the recalls aimed to punish bad judges, civil service reform attempted to decrease the influence of the political machines in the administrative process. Since political machines were used by business to select and coopt judges, eliminating their influence raised the cost of subverting justice, i.e., the $K(A)$ function. Since this also raises X, civil service reform can also be seen as an efficient response to the increased scale of activity.

An alternative to using regulation or raising B was to reduce S directly. In his platform of the New Freedom, Woodrow Wilson opposed “the [Roosevelt] doctrine that monopoly is

⁷ One can argue that when a regulator ultimately needs to go to court to obtain a judgment, the problem of subversion of courts remains. Courts, however, are much less likely to be subverted by defendants when action is brought by organized and powerful regulators, then when the plaintiffs are individual victims.

inevitable and that the only course open to the people of the United States is to submit to and regulate it.” Instead, he advocated reducing firm size “to prevent private monopoly by law, to see to it that the methods by which monopolies have been built up are legally made impossible.”

Indeed, the whole trust-busting program aimed at eliminating the power of trusts to coerce politics and justice as much as reducing prices. Hofstadter (1955, p. 227) writes:

The progressive case against business organization was not confined to economic considerations, nor even to the more intangible sphere of economic morals. Still more widely felt was a fear founded in political realities—the fear that the great business combinations, being the only centers of wealth and power, would be able to lord it over all other interests and thus put an end to traditional democracy

Our model sheds light on such comments. Competitive prices may prevail in a duopoly with two very large firms, but the scale of these enterprises may enable them to subvert justice. Progressive trust-busting should be seen as a response to the subversion of justice, as well as a means of addressing standard problems of monopoly pricing.

Along similar lines, our model might help explain the support for labor unions among many progressives. Traditionally, economists understand labor unions as a means of restricting labor supply and raising wages. Unions, however, can also become a political “countervailing power” to large firms by projecting comparable economic resources (Galbraith 1952). When large firms meet large unions in a political and economic marketplace, the efforts to subvert justice by one meet countervailing efforts by the other. The formation of labor unions parallels the creation of trusts and large industrial firms as a competitive reaction to subversion.

A final reform proposed by many progressives is the complete control of industry by the government. If large firms subvert justice, or corrupt politicians to extract rents, then government ownership may appear attractive. Because the nature of ownership changes a firm’s objective function (as in Hart et al. 1997), government-owned firms may be less likely to corrupt the system to extract rents (Glaeser 2001). Croly in particular favored this response to subversion of

justice: “if the interest of a corporation is so essentially hostile to the public interest ... the logical inference is not a system of semi-official and semi-private management, but a system of exclusively public management.” This type of reform was unpopular in the U.S., perhaps because the many problems of public ownership were appreciated, but European countries moved in that direction. In those countries weaker rule of law meant that less extreme measures could not be relied on to discipline big firms.

Discussion

Of necessity, a simple model like ours cannot deal with many important aspects of the Progressive era. It may be worthwhile, therefore, to point out some of the key omissions and alternative interpretations.

The first question is why the reforms took place at the time they did, and not before or after? This question is especially important in view of our claim that the switch to regulation was efficient, since it is not typically the case that efficiency considerations drive public policies.

In the United States, the increases in scale that warranted an efficient switch to regulation occurred several years, and possibly decades, before the turn of the century. Compared to efficiency, reforms were significantly delayed. One possible explanation of such stability, and of why it was eventually overcome, is that inefficient methods of market regulation did not matter to enough voters to shape their choices, or at least did not matter enough. By the end of the 19th century, a growing number of voters became involved in the commercial and industrial economy, and thus were affected by the subversion of justice. More and more people working for corporations and railroads were exposed to industrial accidents. Increasing numbers of people lived in the cities and suffered from the subversion of city administration. More and more farmers and small businessmen were expropriated by railroads with market power, or by

larger competitors who made special deals with the railroads. “Politically, the rage of the victims counted for very little in 1840, not much in 1860; by 1890, it was a roaring force” (Friedman 1985, p. 476).

While this rage was the dominant factor behind the success of the progressives, at least three other forces mattered. First, in the middle of the 19th century, relations between the North and the South dominated politics. For years after the Civil War, venal Republican candidates could whip up public support by “waving the bloody shirt,” and reminding voters of the North-South issue. The reforming Horace Greeley was soundly beaten at the polls by Ulysses S. Grant—a corrupt President who symbolized Northern dominance. Second, changing technology in publishing facilitated the rise of popular muckraking journals. Free entry in this national industry made it possible for journals such as McClure’s to thrive without the support of the local business community, and indeed to flourish by attacking industry. Finally, a crucial development in U.S. politics was the ascent of Theodore Roosevelt, who ran both his first term and the election of 1904 on the platform of subversion of the judiciary and of restoration of justice through regulation.

The second important question is whether our insistence on the efficiency of regulatory reforms during the progressive era is correct. This question is especially relevant in light of the fact that Progressivism in the United States followed the period of Populism. The simplest alternative theory of regulatory reform is that its objective was redistribution from the haves to the have nots rather than efficiency.

Redistribution certainly played some role in reforms, but efficiency considerations were extremely important. First, many regulatory reforms were supported not just by the working classes, but also by the urban elite, whose interests were championed both by the muckrakers and by the Republicans Theodore Roosevelt and William H. Taft. This was hardly the populist

reform of William Jennings Bryan. In addition, it is important to remember that, unlike the West European countries, the U.S. during this period does not develop a redistributive welfare state. If redistribution was the principal goal of the progressives, there were much more direct ways of reaching it than regulatory reform.

Third, our theory is derived from the assumption that regulators are less subject to subversion than the courts (because the penalties from regulation are lower than those from liability). In the 20th century, the capture of regulators by the industry through revolving door and other policies became important, and dominated professional thinking about regulation after World War II (McCraw 1984, Stigler 1971). Were the theories of superiority of regulation just a figment of progressive imagination?

We do not believe so. We take on face value the progressives' concern about political appointment, intimidation, and corruption of judges. At the turn of the 20th century, these concerns were vastly stronger than those about revolving door policies of professional regulatory commissions. As the U.S. became more of a democratic and orderly society, concerns about corruption and intimidation of judges have naturally receded, and those about regulatory capture came into view. Not surprisingly, and in line with our argument that policies eventually move toward efficiency, the end of the 20th century saw substitution of litigation for regulation.

Indeed, toward the end of the 20th century, the U.S. saw a different kind of potential subversion of justice, where juries awarded enormous damages to plaintiffs suing corporations for harm suffered in industrial and consumer product accidents. In response to such litigation, corporations rather than the public sought protection from regulation, because the resources required to fight liability claims in courts exceeded the potential costs from accepting regulation.

5. Implications for Law Enforcement.

Figure 1 allows us to consider the desirability of alternative modes of law enforcement in different countries and for different activities. We can think of X – the cost of subverting justice – as varying across countries and across activities. Some countries might have highly independent, disciplined and efficient bureaucrats and regulators, who are invulnerable to political pressure and bribes; this is the situation of a very high X . Other countries might have poor, politically vulnerable, and easily corruptible officials, who cannot stand up to the pressure from private parties they are supposed to regulate; this is a situation of low X . We can also think of X varying across activities: a country might have enough bureaucratic prowess to control violence within its borders, but not enough to administer securities or anti-trust laws. Our framework allows us to consider the consequences of such variations. In this discussion, we take X to be exogenous, even though many progressive and other reforms aimed to increase X .

The first, and arguably most important, message of the model is that in situations of extremely low X , the optimal government policy is *to do nothing*. When the administrative capacity of the government is severely limited, and both its judges and regulators are vulnerable to intimidation and corruption, it might be better to accept the existing market failures and externalities than to deal with them through either the administrative or the judicial process. For if a county does attempt to correct market failures, justice will be subverted, and resources will be wasted on subversion without successfully controlling market failures.

This implication of the model is of great significance.⁸ Some economists (e.g., Stiglitz 1989) see market failures as pervasive in the emerging and transition economies, and recommend heavier regulation of economic activity in such economies than in advanced welfare states. Our

⁸ For many further examples of appropriate institutions, see (2001) World Development Report.

model, in contrast, implies that countries operating at low levels of law and order in the first place should institute fewer laws and regulations of economic activity, because their officials cannot administer more without being subverted. A number of examples illustrate this point.

In 1992, under pressure from Western donors, the government of the Russian Federation established an anti-monopoly commission to address the problems of industrial consolidation. It became immediately apparent that the new commission could not stand up to the political power of large enterprises, and it did not even try to regulate their activities. Instead, the commission started to compile lists of small firms, such as bakeries, taking the position that such firms had the potential of abusing their local market power. Small entrepreneurs had to register with the commission, and often to pay bribes just to get off the lists of potential monopolies. The commission did nothing about the real problems of market power, added a level of regulation of small firms, and provided lucrative opportunities for its own employees (Boycko et al. 1995).

This phenomenon is more general. Most countries in the world, including the poorest ones, require many procedures for new firms to begin operating legally. Most of these procedures on paper have market failure justifications: officials check that the new entrepreneurs do not have criminal records, have professional qualifications and bank accounts, obey sanitary restrictions, and so on. These regulations, however, are often subverted through bribes or operations in the unofficial economy (De Soto 1989). In a cross-section, countries with more regulations of entry exhibit higher corruption and larger unofficial economies, but not superior social outcomes that regulation allegedly aims for (Djankov et al. 2002).

The model, then, helps us reconcile the apparent greater extent of market failures in emerging economies stressed by Stiglitz (1989), with the equally apparent failure of regulation. The problem is not that these countries do not have a need for dealing with social harm. Rather,

the problem is that these countries cannot administer the solutions. When neither the courts nor the regulators can resist subversion, the optimal policy is to leave even imperfect markets alone.

The second message of the model is that, with intermediate enforcement capacity, especially in cases of high social damage from market activities, regulation is efficient. When externalities cause large damages, fines necessary to ensure desirable conduct are very high, and therefore liability regimes are especially vulnerable to subversion. Under such circumstances, court enforcement might fail to achieve efficiency, but regulation of inputs stands a better chance. This case is even stronger if, as Landis (1938) maintains, it is more expensive to subvert career regulators than judges. The United States in the early 20th century may in fact have exhibited the conditions under which, for efficiency reasons, regulation was preferred to tort.

Glaeser, Johnson and Shleifer (2001) give another example of the same phenomenon by comparing the regulation of securities markets in Poland and the Czech Republic in the 1990s, at the time when their per capita incomes were roughly comparable to that of the U.S. during the Progressive era. During this period, the Czech government adopted a laissez-faire approach to securities regulation, expecting the judicial system to fill the necessary gaps. Poland, in contrast, adopted strict regulations patterned after the U.S. Securities Acts, and created an independent regulatory commission to enforce them. The result has been the collapse of securities markets in the Czech Republic, as the existing system of law enforcement failed to deal with pervasive fraud in the market. In contrast, the Polish stock market developed rapidly, and there is some evidence that the Securities Commission actually managed to enforce the existing rules.

This analysis also sheds light on a well-known recommendation of Becker (1968) that an optimal penalty system in the world of costly law enforcement should exhibit low probabilities of detection and conviction with very high penalties or fines. Such a strategy combines desirable deterrence effects with cost-savings on enforcement. A large literature explains why we do not

see such law enforcement schemes in reality (Kaplow and Shavell 2000). Leading explanations of limited penalties include risk aversion or bankruptcy constraints of the violators (Polinsky and Shavell 1979) and considerations of fairness (Dickens et al. 1989). We argue, alternatively, that the defendants are likely to spend enormous resources subverting a Beckerian scheme of high penalties, and may well succeed. The Denise Rich example mentioned above illustrates this point. From our perspective, Becker's solution of raising fines has precisely the reverse effect: it only increases the incentives to subvert justice and diminishes the likelihood of good conduct.

The point of this discussion is not that regulation is always the optimal mode of law enforcement. As Stigler (1971) has taught us, regulators can be captured or subverted, just as the judges can, and in many instances doing nothing is preferred to regulation. The point of these examples is rather that, at least in some instances, administrative enforcement of rules is both feasible and preferred to a liability system, and that failure of courts to deal with the problems of market failure is one indication of the possibility that regulation might do better.

Finally, the model implies that societies with the highest costs of subversion should rely on courts rather than regulators. The reason for this is that the liability system – as long as it is not subverted – can achieve first best efficiency, whereas regulation never can. Is this prediction realistic? Advanced economies appear to have both more litigation and more regulation, so it is difficult to say which they have more of, relatively speaking. At the same time, it is striking that, since the 1970s, the United States has followed the path of significant deregulation of many industries, such as gas transmission, trucking, financial services, and airlines. One interpretation of such deregulation is that many issues can now be effectively addressed in courts, and therefore regulation is less efficient than litigation.

In concluding this section, we return to its general point. The law and order conditions in a country are in themselves a crucial determinant of its optimal strategies for regulating

economic activity. Institutions cannot be built without recognizing where the country is in the first place. The finesse with which courts resolve disputes is appropriate for a country that has high levels of law and order. In the extreme contrast, a country beginning with low levels of law and order should tread gingerly in giving officials more power to shape economic life.

6. Conclusion.

In 1960, Ronald Coase posed a crucial problem for economics: what is the optimal strategy of protecting property rights? Many of Coase's followers have interpreted his article as supporting the "free market" idea that a well-functioning market economy, with well-defined property rights, only requires courts to enforce contracts and tort claims. In such a world, regulation is unnecessary. But Coase's reasoning does not necessarily imply the superiority of the judiciary. Efficiency depends on whether courts, or regulators, or neither, work better to address the problems raised by Coase. Different institutions might be most efficient – and most attractive to a libertarian – under different circumstances.

In this paper, we have presented a model of choice of a law enforcement strategy between doing nothing, regulation, and litigation. We showed that the choice depends crucially on the vulnerability of law enforcement in a country to subversion by powerful interests that might be affected. We argued that the model helps explain the rise of regulation of business in the United States between 1887 and 1917, as well as a range of evidence from other countries.

In conclusion, we come back to some issues that the model has swept under the rug. First, as we already indicated, we have taken the cost of subverting justice to be exogenous and considered optimal enforcement policies in its light. In reality, the laws and regulations themselves can raise the costs of subverting justice. They can do so directly by improving enforcement capacity, or indirectly, by introducing rules that make it harder to subvert

enforcement. The more general way to put this point is that law and order does not come from nowhere: it comes from the government introducing laws and orders that are obeyed. A more complete model would endogenize the cost of subversion, and consider how enforcement strategies can themselves raise this cost.

Second, we consider the problem from the point of view of efficiency, and thereby ignore other costs of regulation. Stigler (1971) argued that regulation is generally acquired by the industry to increase its own profits. McChesney (1987) wrote that regulation is introduced by politicians in order to encourage campaign contributions and bribes. Our framework allows for regulatory capture, campaign contributions, and bribes, but sees them as manifestations of subversion of law enforcement, rather than its *raison d'être* to begin with. In the context of the U.S. Progressive era, it is difficult to see the growth of regulation as being a source of increased industry profits, campaign contributions, or bribes, since both the industry and the venile politicians vehemently opposed the new regulations. Nonetheless, a more general view must obviously consider the fact that, in many instances, Stigler's and McChesney's objectives become central to regulatory design. This possibility mediates in favor of doing nothing.

But despite the caveats, we reiterate the general point. The goal of economic institutions is the same across times and places, namely to secure property and to make perpetrators of harmful acts accountable. But even though the goals are constant, which institutions are appropriate for achieving them varies. A key determinant of appropriateness is law and order in the society. In the United States during the Progressive Era, regulation replaced courts as the efficient institution. In other times and places, the appropriate institutions could be the courts, or even nothing. This focus on the efficiency of alternative arrangements can be helpful in advancing the institutional analysis more generally as well.

References

- Adams, Henry (1887), "Relation of the State to Industrial Action," *American Economic Review* 1, 7-85.
- Becker, Gary (1968), "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76, 169-217.
- Becker, Gary and George Stigler (1974) "Law Enforcement, Malfeasance, and the Compensation of Enforcers," *Journal of Legal Studies* 3, 1-19.
- Boycko, Maxim, Andrei Shleifer, Robert Vishny (1995) Privatizing Russia, Cambridge: MIT Press.
- Callow, Alexander B. (1966), The Tweed Ring, New York: Oxford University Press.
- Chandler, Alfred (1977), The Visible Hand, Cambridge: Harvard University Press.
- Coase, Ronald (1960), "The Problem of Social Cost," *Journal of Law and Economics* 3, 1-44.
- Croly, Herbert (1965), The Promise of American Life, Cambridge, MA: Harvard University Press.
- DeSoto, Hernando, (1989) The Other Path, New York: Harper and Row.
- Dickens, William, Lawrence Katz, Kevin Lang, and Lawrence Summers (1989), "Employee Crime and the Monitoring Puzzle," *Journal of Labor Economics* 7, 331-347.
- Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2001), "The Regulation of Entry," forthcoming, *Quarterly Journal of Economics*.
- Fishback, Price V. and Shawn Kantor (2000), A Prelude to the Welfare State: The Origins of Workers' Compensation, Chicago, IL: The University of Chicago Press.
- Friedman, Lawrence M. (1985), A History of American Law, New York: Simon and Schuster.
- Galbraith, John Kenneth (1952), American Capitalism: The Concept of Countervailing Power, Boston, MA: Houghton Mifflin.
- Glaeser, Edward (2001), "Public Ownership in the American City," Harvard University Mimeo.
- Glaeser, Edward and Andrei Shleifer (2001a) "A Reason for Quantity Regulation" *American Economic Review Papers and Proceedings* 91, 431-435.
- Glaeser, Edward and Andrei Shleifer (2001b) "Legal Origins," NBER Working Paper 8272, Cambridge, MA: National Bureau of Economic Research.

Glaeser, Edward, Simon Johnson, and Andrei Shleifer (2001) "Coase Versus the Coasians," *Quarterly Journal of Economics* 116, 853-899.

Goldman, Eric (1947), Rendezvous with Destiny: a History of Modern American Reform, New York: Knopf.

Hart, Oliver, Andrei Shleifer, and Robert Vishny (1997), "The Proper Scope of Government," *Quarterly Journal of Economics* 112, 1127-1161.

Hofstadter, Richard (1955), The Age of Reform, New York: Random House.

Holmes, Oliver Wendell (1881), The Common Law, Boston, MA: Little Brown and Company.

Horwitz, Morton (1992), The Transformation of American Law 1870-1960, Oxford: Oxford University Press.

Josephson, Matthew (1934), The Robber Barons, New York: Harcourt, Brace & Company.

Kaplow, Louis and Steven Shavell (1999), "Economic Analysis of Law," NBER Working Paper 9960, Cambridge, MA: National Bureau of Economic Research.

Landis, James M., (1938) The Administrative Process, New Haven: Yale University Press.

Lloyd, Henry Demarest (1894), Wealth Against Commonwealth, New York: Harper and Brothers.

Lockard, Duane and Walter F. Murphy (1992), Basic Cases in Constitutional Law, Washington: CQ Press.

McChesney, Fred S. (1987), "Rent Extraction and Rent Creation in the Economic Theory of Regulation," *Journal of Legal Studies* 16, 101-118.

McCraw, Thomas K. (1984), Prophets of Regulation, Cambridge, MA: Harvard University Press.

Nagel, Mark (1999), Supplicants, Robber Barons, and Pocket Banks, Ph.D. dissertation, Department of Government, Harvard University.

Polinsky, A. M. and Steven Shavell (1979), "The Optimal Tradeoff Between the Probability and Magnitude of Fines," *American Economic Review* 69, 880-891.

Posner, Richard (1992), Economic Analysis of Law, Boston: Little Brown.

Posner, Richard (1995), "What Do Judges Maximize?" Chapter 3 in *Overcoming Law*, Cambridge: Harvard University Press.

- Schwartz, Gary T. (1981), "Tort Law and the Economy of 19th Century America: A Reinterpretation," *Yale Law Journal* 90, 1717-1775.
- Schwartz, Gary T. (1989), "The Character of Early American Tort Law," *UCLA Law Review* 36, 641-718.
- Shavell, Steven (1984a), "A Model of the Optimal Use of Liability and Safety Regulation," *Rand Journal of Economics* 15, 271-280.
- Shavell, Steven (1984b) "Liability for Harm Versus Regulation of Safety," *Journal of Legal Studies* 13, 357-374.
- Shavell, Steven (1987), Economic Analysis of Accident Law, Cambridge: Harvard University Press.
- Shleifer, Andrei and Robert Vishny (1998) The Grabbing Hand, Cambridge: Harvard University Press.
- Skocpol, Theda (1992) Protecting Soldiers and Mothers, Cambridge: Harvard University Press.
- Steffens, Lincoln, 1906, "The Shame of Minneapolis," reprinted in Arthur and Lila Weinberg, eds The Muckrakers, 2001. Chicago, IL: University of Illinois Press.
- Stigler, George J., 1971: "The Theory of Economic Regulation," *Bell Journal of Economics and Management Science* II, 3-21.
- Stiglitz, Joseph E., 1989, Wither Socialism?, Cambridge: MIT Press.
- Tarbell, Ida M., 1903, "The History of the Standard Oil Company: The Oil War of 1872—How the 'Mother of Trusts' operated" reprinted in Arthur and Lila Weinberg, eds The Muckrakers, 2001. Chicago, IL: University of Illinois Press.
- Wilson, Woodrow (1913), The New Freedom, New York: Doubleday, Page and Company.
- World Bank (2001), World Development Report 2002: Building Institutions for Markets, Washington, D.C.: Oxford University Press.