Antitrust law is only as good as the mechanisms by which it is enforced. Substance and procedure are not distinct bodies, but part of a continuum of legal and institutional rules, practices, and mechanisms working conjunctively to advance consumer welfare and efficiency. It is impossible to understand the substantive rules without understanding the relevant enforcement mechanisms. Judges tend to formulate liability rules with an eye on enforcement mechanisms. For example, judges tend to be skeptical of the ability of lay juries to decide predatory pricing cases, so they formulate deliberately underinclusive liability rules to thin out the number of predation cases reaching trial. Similarly, the Supreme Court has made it hard to plead conspiracy in cartel cases because trial courts have trouble preventing discovery costs from skyrocketing. Evaluating liability rules in a vacuum, without understanding the institutional considerations that motivate judges, might lead to false impressions about the courts’ views of the merits of various competitive practices.

Many of the procedural and enforcement rules that apply to antitrust cases were not designed for antitrust, but are general features of civil or criminal law. Sometimes, mismatches occur between procedure’s generality and antitrust’s specificity. Generic enforcement methods are not always well-suited to the peculiarities of antitrust.

In the US legal system, antitrust enforcement is decentralized and largely uncoordinated. There are two separate federal antitrust enforcement agencies, fifty state attorneys general with enforcement powers, liberal rules for private enforcement, and a treble damages bounty that draws private litigation entrepreneurs into the antitrust litigation market. Antitrust is enforced both civilly and criminally, publicly and privately, prospectively (for injunction) and retrospectively (for damages or other penalties), formally and informally, and administratively and adjudicatively.

Evaluating this crazy quilt of enforcement mechanisms requires defining the goals of antitrust enforcement, which is the subject of the first part of this chapter. The second part asks what forms of public enforcement are best calibrated to achieve these goals. The third part considers two of the leading issues in private enforcement – standing rules and damages.
I. **Enforcement goals**

The goals of antitrust enforcement are bound up with the goals of antitrust law itself. How antitrust is enforced depends substantially on what antitrust law is intended to achieve. For much of the history of US antitrust law, there was debate and disagreement over antitrust law’s goals. The differing views implied widely varying possibilities about the structure of enforcement. Today, there is broad consensus on the goals of antitrust law, which makes possible a broad consensus on the goals and structure of enforcement.

A. **Deterrence, compensation, and any others?**

The modern consensus among economists and antitrust practitioners is that antitrust law should exist primarily to achieve allocative efficiency and to advance consumer welfare. Although these two goals sometimes conflict when it comes to the specification of liability rules, they are generally in harmony when it comes to antitrust’s enforcement goal. Both allocative efficiency and consumer welfare are best served by an enforcement structure that makes the defendant fully internalize the external cost of the violation – the deadweight loss borne by consumers and monopoly transfer from consumers to producers. Such an approach deters anticompetitive behavior by making socially harmful behavior a negative expected value event.

Deterrence is only one of the recognized goals of antitrust enforcement. The Supreme Court has held that compensation of injured parties is an additional goal, although the Court has seemingly made compensation subsidiary to deterrence. From an economic perspective, it is not obvious why compensation should matter at all. Wealth transfers, whether from consumers to producers or from one business to another business, are an external cost of antitrust violations and can decrease social welfare in a variety of subtle ways. However, economic theory cannot predict with great certainty the social welfare consequences of returning overcharges to the victims of the violation. For example, one might think that wealth transfers from consumers to producers would cause a diminution in net social welfare because producers tend to be wealthier than consumers and money begins to bring diminishing marginal utility returns at higher wealth levels. Hence, compensating the consumers would seem to increase social welfare. But the assumption that producers are wealthier than shareholders is far from universalizable. Consider, for example, a cartel among publicly traded yacht manufacturers whose stock is owned in large portion by employees, small investors, and union pension funds. Compensation cannot be justified as a goal of antitrust enforcement on economic terms, although it may have moral or political justifications.
An additional enforcement goal is prevention through *ex ante*, firm-specific control. Instead of disincentivizing anticompetitive behavior (as in the deterrence model), the prevention model involves *ex ante* scrutiny of specific commercial practices by identified actors. Merger control is a leading example of where antitrust works primarily on an *ex ante* approval basis. Instead of punishing firms that have entered into anticompetitive mergers or seeking to break them up after the fact, the Hart-Scott-Rodino Act requires firms that plan to merge to file a notification with the enforcement agencies, enabling the agencies to scrutinize the mergers before they occur. An issue that will be discussed further below is whether it would be preferable to rely more on such an administrative model of antitrust rather than on the adjudicative model that seeks to ascertain and punish past bad acts.

Deterrence and *ex ante* control are the two primary economic goals of antitrust enforcement. Most other goals (in addition to compensation, discussed above) cannot be justified on primarily economic terms. Although political considerations sometimes enter into enforcement decisions, such considerations are largely outside of the jurisdiction of economics.

**B. Overdeterrence and underdeterrence**

In an ideal world, antitrust decision-makers would simply ‘aim to get it right’ and not worry about whether they were tending more toward over-inclusion or underinclusion. But it is unrealistic to expect that bodies of law are free from systematic tilts toward false positives (erroneous findings of liability) or false negatives (erroneous findings of non-liability). For example, free speech law may be oriented toward false negatives. First Amendment law protects a good deal of speech that has little social value because the costs of disallowing socially useful speech are generally thought to be higher than the costs of protecting socially harmful speech. On the other hand, securities regulation may be oriented toward false positives. Publicly traded companies may be required to disclose more than the optimal amount of information – and pay penalties if they do not – because it is thought that the costs of overdisclosure are less than the costs of underdisclosure.

Whether antitrust should err in the direction of overdeterrence or underdeterrence is a question for both antitrust substance and antitrust procedure. Adjudicatory errors may occur in both directions – false positive and false negative – and at both the liability rule-framing level (through underinclusion or overinclusion) and at enforcement level (through factfinder error). A tendency in one direction in substantive rules can be counteracted by a tendency in the opposite direction in procedural rules. For example, a tendency toward false positives at the substantive
level can be counteracted by the framing of procedural rules (such as
evidentiary exclusion rules), stringency in the requirements for expert tes-
timony, or heightened burdens of proof, that make a finding of liability
less probable.12

As noted at the outset, courts tend to frame liability rules in a deliber-
ately underinclusive manner.13 They also tend to frame stringent proce-
dural rules that weed out before trial all but the strongest antitrust cases.
At both the motion to dismiss and summary judgment stages, courts scruti-
nize the economic plausibility of antitrust claims and dismiss those cases
that lack a sufficiently rigorous foundation in economic theory.14 The use
of these procedural screens necessarily strains out some cases that might
be found meritorious if allowed to proceed to discovery or trial. Thus, the
recent tendency in US antitrust law has been to tilt both the procedural
rules and sustentative liability rules toward underinclusion.

There are several possible explanations and justifications for attitudinal
tilts toward false negatives in both liability rules and procedural rules. I
will suggest three possibilities.

First, the costs of false positives tend to be greater than the costs of
false negatives. In an economy characterized by low regulatory entry bar-
riers, a high rate of innovation, and efficient capital markets, privately
acquired market power may be fragile and perpetually contestable – which
makes the need for antitrust intervention comparatively low. This would
suggest that false negatives are likely to cost relatively little. On the other
hand, false positives in antitrust cases may impose costly constraints on
otherwise well-functioning capital and industrial markets.

Second, courts may err in the direction of false negatives over those
facets of the legal system that they control because those aspects of the
legal system that they do not control tilt toward false positives. In particu-
lar, the false-negative orientation of antitrust’s procedural and substan-
tive rules may be explained by judges’ beliefs that jurors tend to err in
the direction of overinclusion or false positives. This tendency may occur
because jurors misunderstand the complex substance of antitrust law and
manifest populist bias against large corporations that use cut-throat –
although not necessarily exclusionary – competitive tactics.15 If jury avoid-
ance explains a least a portion of the judiciary’s false-negative orientation,
one would expect – or hope – to see judges tilting back toward equilibrium
in equitable or administrative actions brought by the government, which
do not entail juries. In fact, we observe relatively little difference in judicial
attitude toward public and private antitrust cases.16

Finally, contemporary judges may be tilting toward false negatives
in reaction to a history of perceived error in the opposite direction. The
Chicago School critique of the interventionist antitrust precedents of
the Warren Court and earlier eras has exerted a profound influence on
the courts. Judicial pendulums sometimes swing to the opposite extreme
before coming to rest in the middle. Antitrust enforcement may presently
be biased toward underinclusion simply because it was formerly biased
toward overinclusion.

II. Public enforcement
US public enforcement is comparatively decentralized. Two different
federal departments or agencies enforce federal antitrust law, as do each
state’s attorney general. The Sherman Act is enforced both criminally and
civilly. On the civil side, the Justice Department can seek both civil penal-
ties and injunctions, and the injunctions may be simple or complex. These
various enforcement mechanisms interact in complex ways.

A. Executive or agency
Both the Antitrust Division of the Department of Justice (and the
regional United States Attorneys offices, which are subsets of the Justice
Department) and the Federal Trade Commission (FTC) enforce the anti-
trust laws. The Justice Department and FTC enjoy concurrent enforce-
ment authority over some statutes and exclusive authority over others.\textsuperscript{17}
However, the two agencies effectively exercise co-extensive authority over
all antitrust (with the exception of criminal enforcement, which is the
exclusive prerogative of the Justice Department).

In theory, one might justify the existence of two federal agencies on the
grounds of comparative advantage over different kinds of matters. The
FTC is set up to be politically independent and technocratic. It enjoys
rule-making powers and can try matters before specialized administra-
tive law judges, rather than generalist Article III judges. Power is dis-
persed among five commissioners, no more than three of whom can be
of the same political party. By contrast, the Department of Justice enjoys
the advantages of unitary executive control, which can accelerate and
streamline decision-making.

Unfortunately, there is very little correspondence between the agen-
cies’ comparative advantages based on institutional structure and their
division of labor.\textsuperscript{18} For example, in 2002 the Antitrust Division and the
FTC entered into a formal clearance agreement in order to avoid duplica-
tion of investigations.\textsuperscript{19} The agreement divided antitrust enforcement
responsibility based on the agencies’ comparative expertise and experience
with different industry sectors, not the institutional structure of the agen-
cies. Thus, for example, the FTC was to investigate computer hardware,
energy, healthcare, retail stores, pharmaceuticals, and professional serv-
ices and the Antitrust Division agriculture, computer software, financial
services, media and entertainment, telecommunications, and travel.\textsuperscript{20} That the Justice Department was to handle computer software while the FTC handled computer hardware had nothing to do with hardware being better suited to the institutional capabilities of the FTC. It was simply a convenient division of labor based on what the two agencies had done in the past. Although the clearance agreement quickly folded due to political pressure from Congress, it exemplifies the essential fungibility of the two agencies.

Not surprisingly, calls have been made to consolidate enforcement in a single agency. For example, this might be accomplished by taking away the FTC’s antitrust enforcement powers and leaving it only a consumer protection/anti-fraud mission. Nonetheless, the institutional status quo seems secure for the foreseeable future. Although very few people would draw up the institutional status quo if working on a blank slate, \textit{tabula rasa} design is a very different question from whether to dismantle a system that, whatever its quirks, seems to be working reasonably well.

\textbf{B. Federal or state}

State attorneys general can enforce federal antitrust law in three ways: (1) as ‘persons’ qualified to seek injunctive relief under Section 16 of the Clayton Act; (2) as persons injured in their business or property when the antitrust violation has harmed the state in its proprietary capacity (\textit{i.e.}, the state government has purchased software from Microsoft); and (3) as \textit{parens patriae} on behalf of their residents.\textsuperscript{21} The states attorneys general can also sue in various capacities to enforce their respective state antitrust laws.

State antitrust enforcers have been perceived as being increasingly active in the last two decades, perhaps in response to less aggressive enforcement in Washington. Some commentators have viewed state enforcers through a public choice lens and accused them of pursuing parochial and localist business interests instead of consumer welfare.\textsuperscript{22} Others have complained that state enforcers have interfered with federal antitrust enforcement. Richard Posner, who attempted to mediate a settlement in the \textit{Microsoft} case, later complained that the participation of the states made it more difficult to coordinate a settlement and interfered with the federal government’s efforts to resolve the matter.\textsuperscript{23} Posner has proposed that the federal enforcers should have the authority to preempt state antitrust enforcement in particular cases.\textsuperscript{24}

Despite such criticisms, there is no doubt that state enforcement of antitrust law can be a valuable complement to federal enforcement, particularly when it is focused on local market conditions over which the states have a comparative advantage. In recent years, the National Association of Attorneys General (NAAG) has made increasing efforts to coordinate...
enforcement among the states, to systematize and regularize state enforce-
ment protocols, and to achieve greater transparency by making publicly
available a database describing the states’ enforcement activities. Some
commentators have viewed state enforcement as a valuable counterweight
to periodic variations in the vigor of federal enforcement, due to changes
in administration.

C. Criminal or civil
Sections 1, 2, and 3 of the Sherman Act, Section 3 of the Robinson-Patman
Act, and Section 14 of the Clayton Act provide for criminal penalties. Yet,
while a wide range of antitrust activity could potentially subject individual
defendants and corporations to criminal fines and (in the case of individu-
als) imprisonment, the Justice Department today prosecutes criminally
only against hard-core cartel behavior such as covert price fixing, bid
rigging, and market division.

Figure 1.1 shows the level of Antitrust Division case filings, adjusted for
the amount of economic activity in the country. The figure shows raw
civil and criminal case filing numbers and antitrust filings as a percent-
age of real GDP, which allows a historical comparison of agency filings
adjusted for overall economic activity. Two aspects of the data are signifi-
cant. First, overall Department of Justice enforcement – at least measured
by case filings – has declined significantly as a percentage of the economy
in the last two decades. Second, the ratio of civil to criminal enforcement
has varied much more historically than the overall ratio of enforcement to
economic activity. Thus, for example, during the Reagan administration
criminal enforcement increased considerably even while civil enforcement
decreased considerably.

The ratio of civil to criminal enforcement depends in large part on the
administration’s enforcement priorities. Criminal enforcement will rise
when the administration views covert behavior, such as price fixing, as a
relatively greater menace than publicly disclosed behavior, such as exclusion-
ary joint venture bylaws. Criminal enforcement is justified by the need
to make covert collusive behavior an \textit{ex ante} negative expected value activ-
ity. If a cartel believes that there is only a 10 per cent chance that it will be
cought, the penalty for being caught must be at least ten times the cartel’s
eXpected profits from the collusion. If it is less, it will make economic sense
for the cartel to proceed. Since the treble damages available in private
cases are probably not enough to make collusion a negative expected value
event (more on this below), some stronger deterrent may be needed.

There are two ways to take the expected profitability out of collusion.
One, just discussed, is to increase the penalty. Another way is to increase
the probability of detection. In the past decade, the Justice Department
Figure 1.1  DOJ cases filed per five-year period

has found a highly effective means of increasing the probability of detection – offering leniency to members of the cartel who disclose the cartel’s existence before it is otherwise detected.28 Such leniency effectively exploits the prisoners’ dilemma facing cartels – sticking together is optimal, cheating first is next best, finding out that another member cheated first is pessimal.

Despite government claims that criminal enforcement is increasing in effectiveness, it is hard to know just how effective anti-cartel enforcement is. Between 1997 and 2006, 156 antitrust defendants were sentenced to incarceration for a total of 64,852 days, an average of 416 days per defendant.29 Thus, the average defendant faces just a little bit more than a year in prison for price fixing. One is tempted to compare the marginal costs and benefit of this expectation of prison to the marginal costs and benefit of adding an additional factor to the damages multiplier – for example, increasing the damages multiplier from three to four for cartels – but the trade-offs between criminal and civil enforcement are never that simple. The individual corporate managers who engage in price fixing may be impervious to further increases in the monetary penalty since their own ability to pay a judgment individually was surpassed long before the multiplier reached three. Hence, criminal liability and civil damages liability may be sending incentives to very different entities – criminal to individual managers and civil to the shareholders, who should respond by engaging in more effective monitoring of their managers.

D. Injunctions and administrative solutions
Antitrust injunctions can take various forms, from short, simple and modest to long and complicated. The simplest form – ‘cease and desist’ orders – require only the defendant to refrain from doing a specified anticompetitive act. Often, however, the enforcement agencies opt for open-ended consent decrees with elaborate protocols for future judicial supervision of the defendant’s behavior. The ‘Paramount decrees’, which impose a variety of complex restrictions on vertical integration and horizontal practices in the movie business, have remained in place (albeit with some relaxations) since 1948.30 In a recent study, Richard Epstein makes a compelling case for less ambitious, less intrusive, and shorter-lasting consent decrees.31 Epstein sensibly argues that consent decrees work best when the decree’s prohibitions are tied directly to the underlying antitrust violations and have a predetermined sun-down, as in the provision terminating the Microsoft consent decree after five years.32

On the other hand, consent decrees may in some circumstances replace antitrust liability with a quasi-contractual and privately enforceable regime that may be cheaper to administer and more effective at regulating
market power than the threat of antitrust liability. A leading example of public–private antitrust is the rate-setting mechanism under the BMI and ASCAP consent decrees, which is now codified in a federal statute. BMI and ASCAP are music performance rights clearing houses that aggregate and license millions of individual artists’ performance rights. The transactions costs of individual licensing negotiations between each artist and each potential licensee make the clearing houses very economically efficient. At the same time, the aggregation of millions of licenses in the hands of large collective bargaining agents creates a substantial amount of market power and the potential for anticompetitive abuse. The solution has been a long-term consent decree resulting from an antitrust action brought by the Justice Department. Under the consent decrees, BMI and ASCAP must make through-to-the-listener licenses available for public performances of their music repertoires and provide applicants with proposed license fees upon request. If the clearing houses and the applicant cannot agree on a fee, either party may apply to the rate court for the determination of a reasonable fee.

It is uncertain whether the rate-setting court solution to the market power problem is effective. There have been relatively few rate-setting proceedings under the BMI and ASCAP consent decrees and virtually none under other rate-setting provisions for intellectual property in antitrust consent decrees. Just as business firms often bargain in the shadow of antitrust law, so too do they bargain in the shadow of rate-setting courts.

Another form of antitrust enforcement that is largely informal and administrative in character is merger review, under the Hart-Scott-Rodino Antitrust Improvements Act of 1976. Hart-Scott specifies that, as to certain classes of stock and asset acquisitions, the acquiring and/or acquired person must file a premerger notification with the Department of Justice and FTC. Unless the agencies give early termination, the merger or acquisition cannot close for 30 days following the filing. Prior to the termination of the 30-day waiting period, the agencies can issue a ‘second request’, a species of subpoena for categories of documents and information additional to those that must automatically accompany the initial filing. The agencies can then extend the waiting period for 30 days following satisfaction of the second request. Formally, compliance with Hart-Scott does not mean that a merger is approved or that the merger is deemed legal. But the effect of Hart-Scott has been to create a de facto administrative regime of merger approval by government economists and antitrust lawyers who consider, ex ante, the likely structural consequences of a merger and negotiate with the merging parties for divestiture packages or conduct commitments sufficient to alleviate competitive concerns. Merger practice has become an administrative enterprise conducted by
federal industrial policy experts with wide powers to specify the structure and competitive behavior of merging corporations.45

Administrative solutions have many potential advantages over conventional adjudication in furthering antitrust values. Conventional adjudication is largely binary – i.e., the merger is lawful or it is not; the defendant did or did not monopolize. Administrative processes can come up with more fine-tuned solutions to the problem of market power. Conventional adjudication tends to delegate decision-making to generalist judges and lay jurors. Administrative solutions tend to be more technocratic and involve decisions by experts. Conventional adjudication tends to be backward-looking (damages, deterrence) while administrative solutions are often forward-looking (rate-setting, merger-structuring).

It is conventional to juxtapose antitrust adjudication and regulation as competing modalities of economic control, but administrative solutions need not be conventionally regulatory or entail centralized command-and-control regulation. As noted, enforcement of the BMI and ASCAP consent decrees is initiated privately. Similarly, in recent years a number of patent pools adjacent to standard-setting organizations have created private administrative mechanisms to set patent royalty rates and other licensing terms in an effort to replace antitrust litigation with a quasi-contractual solution to the problem of market power.46

III. Private enforcement
For every antitrust case filed by the US government (whether the Department of Justice or the FTC), there are approximately ten private cases filed in the federal courts.47 The United States is unique in this regard. In most other jurisdictions with serious antitrust laws, public enforcement is the norm and private enforcement the rare exception. Given the volume of private cases, two enforcement issues become critical: who can sue and how much can they recover?

A. Standing rules
Although antitrust law exists supposedly for the benefit of consumers, consumers do not make up a majority of the plaintiffs who file private antitrust cases. The Georgetown Study of Private Antitrust Litigation, conducted on a sample of 2,500 antitrust cases from 1973–1983, found that one-third of private plaintiffs were defendants’ competitors, another 30 per cent were dealers or distributors, and less than 20 per cent were customers or otherwise consumers.48 The high number of suits by competitors and other business interests is worrisome. Antitrust lawsuits are themselves powerful vehicles for raising rivals’ costs and excluding competition.49

One solution would be to bar competitor suits and limit standing to
injured consumers. But that solution has its own problems. First, some anticompetitive violations never succeed in harming consumers because the defendant fails to achieve monopoly power. Yet, there is much sense in allowing a claim for attempted monopolization and not only the completed act. Second, the injury to consumers is often too diffuse to make consumer suits cost effective. Each purchaser may have only pennies at stake, while the monopoly gains to the defendant, and losses to its rivals and other vertically related businesses, are enormous. Class action treatment, which has its own problems, provides only a partial solution. Rivals of the defendant and other business interests may have informational advantages over consumers in identifying and fighting anticompetitive conduct. Consumers may be unaware of how a dominant firm’s conduct is keeping new competitors from coming to market but the potential new competitors will know.

So, if private litigation is going to remain an integral part of the enforcement system, it is probably not wise to limit standing to consumers. There are other ways to limit abusive suits by rivals or other disadvantaged business interests. I will mention two of them briefly. First, the Supreme Court has vigorously pressed an ‘antitrust injury’ doctrine which requires a plaintiff to show not merely that the defendant committed an antitrust violation but also that the harm to the plaintiff was of the kind with which the antitrust laws are concerned. Thus, for example, in Brunswick Corp v. Pueblo Bowl-O-Mat, Inc. the Court confronted a claim by a bowling center operator who alleged that it was anticompetitive for a bowling equipment maker to integrate vertically and acquire a bowling center chain that was otherwise going into bankruptcy. The plaintiff’s alleged injury was based on the fact that competition continued when, but for the allegedly anticompetitive acquisition, competition would have diminished. Thus, the injury was not the kind that antitrust law was intended to prevent, even if the acquisition itself was anticompetitive. This ‘antitrust injury’ rule has facilitated the dismissal of competitor suits that raise hypothetical antitrust violations but have not resulted in real consumer harm.

A second antitrust doctrine that has weeded out a number of lawsuits is the ‘direct injury’ rule. In a moment, we shall consider this rule in the context of claims by purchasers (the ‘direct purchaser’ issue), but the rule is also invoked to limit suits by rivals and other business interests. The rule is similar to the proximate cause rule of tort law, although it adds some extra wrinkles. The basic intuition is that antitrust violations often cause injury in a falling domino pattern. The defendant organizes a boycott of its rival which harms the rival, the rival’s shareholders (and the shareholders of the rival’s shareholders, and so on), and the rival’s suppliers (and the suppliers of the rival’s suppliers, and so on). All of these actors may
be able to say that they were injured by the antitrust violation, but not all of them should be able to sue. For one thing, there would be a good deal of duplication of damages and windfalls if both the injured rival and its shareholders could sue.\textsuperscript{54} This is less obviously true of the supplier, but if we allowed the rival’s direct supplier to sue, then why not allow the supplier’s supplier to sue, and so forth up and down the economic chain? The direct injury rule tries to cut off standing at the most immediate level of harm, which is often the rival firm. The strong intuition is that the most direct victims of the antitrust violation will also be the best motivated and informed parties to file the antitrust suit and more than capable of performing the deterrence function. Even if some damages from the violation are not incurred by the direct victims and hence not recoverable from the defendant, the automatic trebling should more than make up for any such slippage.

Similar ‘directness’ issues are raised with respect to customer standing to sue. Suppose that a price-fixing cartel raises prices $100 above the competitive level. Ultimately, buyers of the defendants’ product will pay $100 more, but which buyers? If the defendants are manufacturers, the goods may be sold first to a wholesaler, then to a retailer, and then to an initial consumer, and then resold (used, but still reflecting a monopoly mark-up) to a second consumer. If each of these purchasers sued for the full overcharge paid by that person, the recoverable damages would total far more than $100.

The US Supreme Court addressed these concerns in a trilogy of cases. In the first case, \textit{Hanover Shoe Inc. v. United Shoe Machinery Corp.},\textsuperscript{55} the Court held that the direct purchaser (the wholesaler in our example who buys directly from the price-fixer) has standing to sue for the full amount of the overcharge to him, even though he may have passed on the overcharge downstream and suffered no economic damage as a result. In the second case, \textit{Illinois Brick Co. v. Illinois},\textsuperscript{56} the Court held, conversely, that an indirect purchaser (the retailer or customer in our example) cannot sue even though he may have been the party that actually absorbed the overcharge and suffered economic harm. In the trilogy’s third case, \textit{California v. ARC America Corp.},\textsuperscript{57} the Court held that federal antitrust law does not preempt state antitrust laws that allow indirect purchaser suits.

These three cases, while not illogical individually, have made quite a mess of things. First, it is not always easy to determine who is or isn’t a ‘direct’ purchaser so the direct purchaser rule’s chief justification – ease of administration – is often eroded by the creation of exceptions to the rule and extensive litigation over its meaning and application. Second, even if the simplicity and symmetry of the \textit{Hanover Shoe–Illinois Brick} regime has efficiency and deterrence justifications, it also creates the morally
unappealing result of economically uninjured large businesses reaping windfall damages recoveries while economically injured consumers take nothing. A number of states have reacted by effectively repealing *Illinois Brick* (either judicially or legislatively) under their own antitrust statutes and allowing indirect purchasers to sue. But this only creates more havoc, since there are now *Illinois Brick* repealer states, non-repealer states, and states somewhere between. This leads to extreme complexity, choice of law gamesmanship, and forum shopping in antitrust cases concerning national markets.

The congressionally appointed Antitrust Modernization Commission recently made a recommendation for legislative reforms that would overrule both *Hanover Shoe* and *Illinois Brick* and allow for removal of state cases to federal court and consolidation of all damages claims as to a particular violation. The court would then make a determination of what the total monopoly overcharge was, treble the overcharge, and allocate the damages pot to the different plaintiffs based on the proportion of their individual injuries to the total. While this system would entail its own complications, it would provide a strong improvement over the status quo.

### B. Damages rules

From a deterrence perspective, the goal of antitrust damages is to make antitrust violations a negative expected value event and, hence, to discourage anyone from committing an antitrust violation. This much private antitrust enforcement shares with public antitrust enforcement. Defendants are relatively (although not completely) indifferent to the payee of their penalty – whether it be the government or a private party. Thus, an increase in the amount of public penalties can offset a decrease in the amount of private penalties, and vice versa.

As noted earlier, the probability of detection is a crucial input into ascertaining the optimal penalty. If the penalty were set at just the social cost of the violation – roughly, the overcharge from consumers to the defendant, the deadweight costs of forgone transactions, and the costs of enforcement – then there would be suboptimal deterrence, because violations might remain positive expected value events. Thus, the optimal penalty, including both private damages and government fines, is equal to the monopoly overcharge, plus the wealth transfer from consumers to the defendant, plus the costs of enforcement, times the probability of detection. For example, if the social cost of the violation is $100 and it was 20 per cent likely that the violation was going to be detected, the optimal penalty is $500.

Under US antitrust law, private damages recoveries are automatically trebled (juries are not told about this, so unless one of the jurors knows
independently about trebling, it is unlikely that the jury will discount the damages award knowing that it will be trebled). The trebling rule could be justified partially by the fact that not all of the social cost of the antitrust violation is recoverable as damages. In a cartel case, for example, the usual plaintiffs will be purchasers of the price-fixed good or service who paid more as a result of the conspiracy. But those plaintiffs’ loss represents merely the wealth transfer consequence of the violation. Consumers who considered purchasing the defendants’ goods or service but found the price too high and therefore substituted to some second-best solution are the core victims – their injury is the inefficient deadweight loss. But it is very hard to make plaintiffs out of people who did not purchase the defendants’ goods. There are no transactions to be identified and the claim ‘I would have purchased’ is often highly speculative. So most purchaser–victim classes consist of plaintiffs who did transact and paid a higher price, not of the core antitrust victims who are usually unidentifiable.

The trebling rule could reflect a rough intuition that only one of out every three antitrust violations is detected and that most are not publicly prosecuted in any event. Even assuming that this intuition is correct on average, it is very unlikely that probability of detection in antitrust cases clusters toward the mean. To the contrary, there appear to be classes of antitrust cases where detection is highly likely and other classes where it is highly unlikely. According to one study, a cartel’s probability of detection is between 13 and 17 per cent.61 Although the estimates vary considerably, most put the probability of detection below 20 per cent.62 On the other hand, certain types of predation strategies rely heavily on signaling long-term predatory commitments to rivals, and thus are only likely to work if they are detected.63 Some anticompetitive schemes work only by stealth, others only by loud announcement, and yet the undifferentiated treble damages multiplier treats them all as if they operated by the same degree of stealth.

One potential solution is to tailor the damages multiplier to the degree of the concealment. For example, a jury might be asked an initial binary question – did the defendant conceal its anticompetitive behavior – and then, if the answer is yes, make a further decision as to what number – say 25, 50, or 75 per cent – is closest to the likelihood of detection.64 The judge would then multiply the actual damages award by an amount corresponding to the number selected sufficient to make the antitrust violation a negative expected value event.65 Whether this would improve over the status quo is subject to some doubt – introducing more complexity into already complex antitrust trials might just increase the overall error rate.

Although trebling receives the lion’s share of attention on the question
of private remedy, an equally important question concerns what to do with uncertainty about the amount of damages. Antitrust violations disrupt markets and recovering the ‘but for’ world is highly problematic. This problem is particularly acute in claims by would-be new entrants that were excluded from the market by the defendant’s anticompetitive conduct. At common law, plaintiffs who were denied a new business opportunity by some wrongful act of the defendant – say a breach of contract or tort – faced denial of their damages claim under a ‘new business rule’ that denied lost profits to firms that did not have an established track record.\(^6^6\) Even where the ‘new business rule’ was not applied in rule-like form, the plaintiff still bore the burden of proving its lost profits with reasonable certainty. This effectively meant that the costs of the uncertainty created by the defendant’s wrong were borne by the injured party rather than the wrongdoer.

Antitrust law treats the frustrated new entrant’s claim quite differently. As a threshold matter, the plaintiff must prove that it was ‘prepared’ to enter the market – that it had the intention and capability of entering and that it took material, affirmative steps toward entry.\(^6^7\) But once the plaintiff establishes standing, the law effectively shifts the costs of uncertainty about the amount of damages to the defendant.\(^6^8\) In a case where there was damage but the amount is quite speculative, the plaintiff is given considerable leeway in creating a model of lost profits.

In combination leniency in proof of the amount of damages and trebling create the possibility of overdeterrence – that is to say, that the law will deter socially beneficial conduct. This is particularly a concern given that many of the cases where damages are most speculative are lost profits claims by allegedly foreclosed new entrants – cases where, unlike cartel cases, the probability of detection is very high because the harm is concentrated in a single entity and the conduct is visible.

There is no formulaic solution to the problem of uncertainty over damages awards. Verbal formulations – ‘reasonable certainty’, ‘malfeasant should bear the costs of the uncertainty’, and so on – fail to provide meaningful guidelines for courts. Perhaps the best that can be suggested is that courts should play a rigorous gatekeeping role on expert testimony about damages, ensuring that damages estimates are based on reliable benchmarks and credible economic theories.

Notes
1. Professor of Law, University of Michigan.

5. See Hovenkamp, supra n. 4 at 1. The reasons that antitrust actually exist may be quite different. As George Stigler has pointed out, the actual purposes of antitrust can only be derived from its effects. Stigler, G. (1975) ‘Supplementary Note on Economic Theories of Regulation’, in *The Citizen and the State*, Chicago, IL and London, UK: University of Chicago Press. In many cases, the effect of antitrust enforcement has been to shift wealth to politically advantaged interest groups rather than to advance the welfare of consumers. See generally McChesney, F. & W. Shughart (eds) (1995), *The Causes and Consequences of Antitrust: The Public Choice Perspective*, Chicago, IL and London, UK: University of Chicago Press.


7. Divergences between total welfare and consumer welfare standards may be more significant in theory than in practice. Tom Barnett, who is currently the Assistant Attorney General in charge of the Antitrust Division, reports that from his perspective as an enforcer, ‘the consumer welfare and total welfare standards can diverge, although I think it is a rare case in practice’. Barnett, T. (2005), ‘Substantial Lessening of Competition – The Section 7 Standard’, *Colum. Bus. L. Rev.*, 2005, 293, 297.


9. *Illinois Brick Co. v. Illinois*, 431 U.S. 720, 746 (1977). In *Illinois Brick*, the Court held that only ‘direct purchasers’ have standing to sue for overcharges resulting from anti-competitive behavior, a holding that may be satisfactory from a deterrence perspective but not from a compensation perspective since the ‘direct purchasers’ are often wholesalers or retailers who simply pass on the overcharge to the consumer.


Antitrust law and economics

15. See Crane, Technocracy, supra n. 11 at 1217.
16. Id. at 1187–8.
17. In general, the Department of Justice enforces the Sherman Act, the FTC enforces the FTC Act (which means basically the same thing as the Sherman Act), and both agencies enforce the Clayton and Robinson-Patman Acts.
18. Some of the FTC’s potential comparative advantages are not exploited. For example, the FTC does not engage in rule-making on issues of antitrust substance even though a statute gives it that power.
20. Id. at Appendix A.
26. The underlying data are pooled from two sources. For 1938–2006, the data were provided to the author by the Department of Justice Antitrust Division. For earlier years, the data were drawn from Posner (1970), ‘A Statistical Study of Antitrust Enforcement’, at 82, Table 5.5. Posner’s data were compiled from the Commerce Clearing House (‘CCH’). For the years in which there was overlap between Posner’s CCH data and the data provided by the Department of Justice, there were slight but relatively insignificant differences. Real GDP numbers are chained 2000. Numbers from 1929 forward are collected from Bureau of Economic Analysis, http://www.bea.gov/national/xls/gdplev.xls. Pre-1929 numbers are collected from http://eh.net/hmit/gdp/.
27. A fuller discussion of these data and their meaning appears in Crane, Technocracy, supra n. 11 at 1175–6.
31. Id.
32. Id. at 112–15. But see Hovenkamp, H. (2005), The Antitrust Enterprise: Principle and
The economics of antitrust enforcement


35. Id.

36. Broadcast Music, 426 F.3d at 95.


40. Id. at § 18a(b)(1).

41. Id. at § 18a(e).

42. Id. at § 18a(e)(2).

43. Nor does the agency’s negotiation of a divestiture package preclude private parties from arguing that the agencies did not go far enough to ensure a competitive market. See Six West Retail Acquisition, Inc. v. Sony Theatre Management Corp., No. 97 CIV. 5499(DNE), 2000 WL 264295, at *23 (S.D.N.Y. Mar. 09, 2000).


45. See Crane, Antitrust Antifederalism, supra n. 2 at 52–3.


47. See Crane, Technocracy, supra n. 11 at 1178, 1182.


50. The classic articulation of the existence of a separate category of attempted monopolization is Justice Holmes’s opinion in Swift & Co. v. United States, 196 U.S. 375 (1905).


53. The leading antitrust case in this area is Associated Gen. Contractors of Cal., Inc. v. Carpenters, 459 U.S. 519 (1983), where the Court denied standing to union that alleged that a contractor’s association had coerced third parties to do business with non-unionized labor suppliers. The Supreme Court has reinforced the direct injury rule in several more recent Racketeering Influenced Corrupt Organizations Act (RICO) decisions. See Anza v. Ideal Steel Supply Corp., 547 U.S. 451 (2006); Holmes v. Securities Investor Protection Corporation, 503 U.S. 258 (1992).

54. It is also the case that some shareholders will be made whole by the rival’s damages recovery, although the present shareholders may not be ones who suffered the original loss and the securities markets may not have fully valued the rival’s antitrust lawsuit.


58. Report and Recommendation of the Antitrust Modernization Commission Chapter

59. The reason for the ‘although not completely’ caveat is that firms would rather not pay their competitors damages, since such payments entail not merely reducing the defendant’s own funds but also (potentially) weakening its market position as a competitor is strengthened.

60. See Hylton, supra n. 8 at 43–7.


64. It would seem silly to include the number zero, since the conduct necessarily was detected if it is now in litigation.

65. For example, if the violation were only 25 per cent likely to be detected, it would be necessary to quadruple the actual damages – plus a little bit – to make the violation a negative expected value activity.

66. See, e.g., MindGames, Inc. v. Western Publ’g Co., 218 F.3d 652, 656–7 (7th Cir. 2000) (explaining common law rule).

67. E.g., Ashley Creek Phosphate Co. v. Chevron USA, Inc., 315 F.3d 1245, 1254–5 (10th Cir. 2003).

68. Bigelow v. RKO Radio Pictures, 327 U.S. 251, 265–6 (1946) (the most elementary conceptions of justice and public policy require that the wrongdoer shall bear the risk of the uncertainty which his own wrong has created).

References


DeBow, M. ‘State Antitrust Enforcement: Empirical Evidence and a Modest Reform
The economics of antitrust enforcement


22 Antitrust law and economics


Cases
Ashley Creek Phosphate Co. v. Chevron USA, Inc., 315 F.3d 1245 (10th Cir. 2003).
MindGames, Inc. v. Western Publ’g Co., 218 F.3d 652 (7th Cir. 2000).

Statutes