Norms in Law and Economics

First draft: April 7, 2003
This draft: March 29, 2005

Richard H. McAdams and Eric B. Rasmusen

Abstract
Everyone realizes the importance of social norms as guides to behavior and substitutes for law, but coming up with a paradigm for analyzing norms has been surprisingly difficult, as has systematic empirical study. In this chapter of the Handbook of Law and Economics, edited by A. Mitchell Polinsky and Steven Shavell and forthcoming in 2006, we survey the topic.

McAdams: Guy Raymond Jones Professor, University of Illinois College of Law, 504 East Pennsylvania Avenue, Champaign, Illinois 61820. Office: (217) 333-4385. Fax: (217) 244-1478. Rmcaadams@law.uiuc.edu.


Copies of this article can be found at http://www.rasmusen.org/papers/norms.wpd. We thank Jonathan Baron, Lisa Bernstein, F.H. Buckley, Robert Ellickson, Timur Kuran, Thomas Miceli, Geoffrey Miller, Peter Ordeshook, Richard Posner, J. Mark Ramseyer, Barak Richman, and Thomas Ulen for their comments, and Sean Mead for research assistance.
1. Introduction

Law seeks to regulate behavior when self-interest does not produce the correct results as measured by efficiency or fairness. If people behave well without regulation, law is superfluous and just creates extra costs. If law is not what actually determines human behavior, scholars debating it are wasting their time. For this reason, law matters primarily to the “bad man” of Oliver Wendell Holmes, Jr. (1897). The “bad man” is, in effect, “economic man,” caring only about the material consequences of his actions:

You can see very plainly that a bad man has as much reason as a good one for wishing to avoid an encounter with the public force, and therefore you can see the practical importance of the distinction between morality and law. A man who cares nothing for an ethical rule which is believed and practised by his neighbors is likely nevertheless to care a good deal to avoid being made to pay money, and will want to keep out of jail if he can.

The man who is not “bad” in this sense, however, is influenced by the ethical rule, either because he cares directly about it or because he cares about other people who do. Since the perfect “bad man” is untypical, we should revise our first sentence above to say that law becomes relevant only when neither self-interest nor social norms provide adequate incentives for behavior.

Since the early 1990s, considerable scholarship in law and economics has turned its attention to norms, as Ellickson (1998) details. Numerous articles and at least six law review symposium issues have addressed the power of social norms and their relevance to law (see “Symposium...” in the References section). Something else Holmes said is: “For the rational study of the law the blackletter man may be the man of the present, but the man of the future is the man of statistics and the master of economics.” And indeed, the same economic methods useful for analyzing law are useful for analyzing norms, a tradition going back as far as Adam Smith (e.g., his explanation in The Wealth of Nations (1776, Book V, Chapter 1) of how religious sects flourish in the anonymity of cities to provide indicators of good morals). Economics is eminently suitable for addressing questions of the various incentives mediated neither by the explicit price of some good nor by the threats of government, incentives that underlie norms such as guilt, pride, esteem and disapproval.

We will proceed as follows. Section 2 addresses the definition of “norms” and contrasts it with “conventions.” Section 3 discusses the sources and workings of conventions and norms, paying particular attention to the normative incentives of guilt and esteem. Section 4 provides a general overview of the norms literature in law and economics, separately discussing how such regularities matter to the positive and
normative analysis. Section 5 reviews applications of this literature to particular areas of law—torts, criminal law, constitutional law, and so forth. Section 6 concludes.

2. Defining “Norms”

Ellickson’s seminal work, *Order Without Law* (1992: 126) notes a fundamental ambiguity in the word *norm*, that it denotes “both behavior that is normal, and behavior that people should mimic to avoid being punished.” Confusion arises because law and economics scholars use the term in both senses. All contributors to the literature seem to agree that a norm at least includes the element of a behavioral regularity in a group—what is typical or “normal”—but they do not agree on whether a norm also requires that the behavior be normatively required. “Norm” means merely equilibrium behavior in Picker (1997); Mahoney & Sanchirico (2001, 2003); and E. Posner (2000). Others, however, restrict the term to the combination of an attitudinal regularity and a behavioral regularity—i.e., the situation where people believe that the behavior is normatively appropriate (Cooter 1996; Ellickson 1992; Kaplow & Shavell 2001a, 2002a, 2002b; McAdams 1997, 2001).1 The attendant attitude may be as strong as a perceived moral obligation—that most people believe that everyone should conform to the regularity and that it is wrong to do otherwise (Cooter 1996; Kaplow & Shavell 2001a) – or as weak as a simple sense of approval or disapproval (McAdams 1997; Pettit 1990). Normative attitudes not only add a distinct element to a behavioral regularity, they also contribute to stability by creating the *normative incentives* – guilt, esteem, shame – that we discuss below.2

Here we will define “norms” as behavioral regularities supported at least in part by normative attitudes. We will refer to behavioral regularities that lack such normative attitudes as “conventions.” This is because we think it useful to have one term—“convention”—for a mere equilibrium that plays out without anyone holding beliefs about the morality of the behavior, and another term—“norm”—for a behavioral regularity associated with a feeling of obligation. This usage also aligns with that in other

1 We include Ellickson (1992:124), whom we read as implicitly referring to normative attitudes when he describes norms as a form of “social control,” where “social control” means enforced rules of “normatively appropriate behavior.”

2 Ellickson notes (1992: 128) that “the best, and always sufficient, evidence that a rule is operative is the routine ... administration of sanctions ... upon people detected breaking the rule.” Although we agree that third-party sanctions commonly reflect the existence of an attitudinal pattern – that the third parties believe the sanctioned behavior violates an obligation or at least that they disapprove of it – game theory shows that such an attitudinal pattern is not strictly necessary. See Mahoney & Sanchirico (2003). Third party enforcement can, in theory, exist merely as a matter of convention.
social sciences. By contrast, if norms are nothing but behavioral regularities without support from attitudes, norms are not really a subject distinct from game theory. Indeed, the concept of “norms” under the broad definition has been justly criticized by such scholars as Kahan (2001) and Scott (2000) as too broad to be useful.

In excluding conventions, we clearly exclude some of what the law-and-economics literature has discussed as “norms” – for example, the equilibria that emerge from the evolutionary models of Picker (1997) and Mahoney & Sanchirico (2001; 2003), and the signalling model of Eric Posner (2000). Similarly, we exclude what Hetcher (2003) calls “epistemic norms,” regularities that arise when individuals faced with information scarcity follow the crowd as in the cascades of Banerjee (1992) and Bikhchandani, Hirshleifer & Welch (1992). All these contributions are useful, but we see their point as explaining what seem to be norms, motivated by feelings of right and wrong, as really being something else – conventions motivated by simple self-interest.

Even using the narrow definition of norm, conventions remain relevant. First, conventions are invaluable for testing whether a norm-based explanation is strictly necessary. As a first step, ask of each behavioral regularity whether it is really due to a convention. Often it will be, and there is no need to employ the special tools of this chapter. Second, conventions sometimes explain the origin of norms. Human beings quickly come to hold normative attitudes about an existing state of affairs, believing that other people should do what they are expected to do, especially when unexpected behavior causes harm (Sugden [1998]). Once everyone expects motorists to drive on the right side of the road, we come to believe that someone who drives on the left is not just foolish, but immoral. What is at first merely a convention becomes a norm. In such cases, an understanding of what maintains the end state requires the idea of norms, but the best tools for understanding norm origin come from game theory. We will therefore discuss conventions in some detail below.

Aside from definitions, there remain other sources of confusion that we hope to avoid. First, although sociologists and anthropologists refer to “legal norms,” we will, following the convention of the legal literature, discuss norms as distinct from law. Although we comment below on the two important meta-law norms of legal obedience and the rule of law, we view law and norms as distinct incentives for behavior. Second, some theorists use “norms” to refer only to decentralized and informally created regularities, while others use the term to refer to rules of private institutions or organizations – rules that are often highly centralized and formal. We consider norms to encompass both types of regularities, though we recommend the term “organizational norms” to refer to centralized norms. Third, scholars such as Miller (2003) and Strahilevitz (2000, 2003) refer to norms that arise between strangers in large populations, whereas others, such as Bernstein (1992) and Ellickson (1992), discuss the norms of
small and close-knit subpopulations. Norms in the sense we study here arise in both settings, though we will use the term “group norm” to refer to norms limited to a particular group. Finally, some theorists implicitly reserve the term “norms” to refer only to general regularities, such as the norms of reciprocity or individualism, while others use the term for specific regularities, such as giving gifts on Secretary’s Day or shutting off cell phones in church. Norms under our definition encompass regularities at all levels of generality.

It is also important to distinguish norms from the rules of thumb and psychological heuristics studied by behavioral economics. Books such as Kahneman, Slovic & Tversky’s 1982 *Judgment Under Uncertainty: Heuristics and Biases* and Dawes’s 1988 *Rational Choice in an Uncertain World* document and discuss many cognitive biases and compensating heuristics, but it is quite possible for a decisionmaker to be perfectly rational yet driven by norms, or radically irrational yet indifferent to norms. If most individuals in a social group eat spinach ice cream, a conventional economist might rest content with the explanation that they like the flavor, while a behavioral economist might attribute that odd behavior to a bias or heuristic. A norms scholar, in contrast, would look for whether there was a desire to conform to what others expect and approve and would check to see if people in the group believed eating spinach ice cream was morally obligatory. Heuristics and rules of thumb do have important implications for laws and lawmaking (see, e.g., Baron (2001)), and they have been called norms (e.g., Epstein [2001]), but they really are a different subject. Psychology does, however, have application in the experimental study of what people mean by such things as “fairness,” as may be seen in Thibaut and Walker’s 1975 *Procedural Justice: A Psychological Analysis* and the literature that followed it (e.g., the criticism in Hayden & Anderson (1979) and Rabin (1993)). Much may be discovered by experiments such as those of Cox & Deck (2004) that do not investigate only whether people behave as the simplest economic models of rational and selfish decisionmaking predict, but also carefully distinguish between different possible motives for deviating from the simple model.

3. How Norms Work

In this section, we will first discuss what we mean by normative incentives, and then contrast that with the numerous ways in which conventions can imitate, generate, or sustain norms.

a. Types of Normative Incentives

People feel obligations in a variety of ways, some internal and some external. Normative incentives are frequently negative – costs imposed on those who fail to
conform to a behavioral regularity (such as guilt from not protecting a child from drowning) – but can also be positive – benefits conferred on those who exceed the normative requirement (as a person who incurs great danger saving a stranger’s life). A significant literature documents and discusses negative sanctions, usually imposed by third parties but sometimes by the victim of a norm violation. Examples include gossip (Ellickson 1992:214-15; McAdams 1996); admonishment and insult (Miller 2003; Buckley 2003); social ostracism and shunning (E. Posner 1996a); economic boycott and exclusion (Bernstein 1992, 1996, 1999, 2001; Skeel 2001); property destruction (Ellickson 1992: 215-19; Miller 2003: 931); and violence (McAdams 1995; Milhaupt & West 2000: 68). Positive sanctions have received less attention, but see Ellickson’s (1991: 236-39) discussion of rewards for third party norm enforcers.

Whether they are positive or negative, by “normative incentives,” we do not mean merely these external sanctions, which are the proximate but not ultimate influence on behavior. Instead, we must ask why third parties ever bother to incur the costs of sanctioning norm violators. Often it is possible to explain the third party behavior as itself part of a convention, not dependent on normative beliefs– see Mahoney & Sanchirico (2003) or Hirshleifer & Rasmusen (1989). Underlying a norm in the strict sense of the word, however, is a non-material motivation, either for the primary behavior of the person who follows the norm or for the secondary behavior of the people who reward his conformity or punish his violation.

The place to look for norms as opposed to conventions, therefore, is in the utility function. Normative attitudes are beliefs about the appropriateness of behavior, and the starting point for analysis is how these beliefs influence utility. Consider three possibilities.

(i) Guilt and pride. An internalized normative incentive means that an individual sanctions himself. Guilt is disutility that arises when a person behaves in ways he thinks morally wrong. The converse, pride, is utility that arises when he behaves in ways he thinks virtuous. That someone can feel guilt and pride is equivalent to saying that he has a taste for behaving in conformity with his moral beliefs. Moral philosophers have for a considerable time emphasized the role of guilt and pride in moral behavior, as in Hume (1751:150). These incentives do not require that anyone else know how the person acted. Nor do they preclude the individual from acting contrary to his moral beliefs – sometimes the payoffs for doing so are greater than the anticipated guilt costs. As elsewhere in economics, in this style of analysis the individual calculates what maximizes his utility and acts accordingly. As elsewhere, the empirical prediction is that when prices change, so will behavior: if the material benefit of norm violation rises while the guilt penalty stays constant, we will observe more violation.
As with other tastes in the utility function, if a person’s taste for pride and distaste for guilt varies widely from day to day, the rational-actor approach will not yield useful predictions. By the end of childhood, however, the moral beliefs that underlie guilt and pride are fixed enough to be difficult to change. Some psychologists claim that there is a genetic basis for guilt. This idea has been picked up in law-and-economics by Rubin (1982), Richard Posner (1986) and Kaplow & Shavell’s *Fairness and Welfare* (2002) because evolutionary theory can explain moral tastes in the same way that it explains the taste for leisure or sweets. Part of the evolutionary explanation is the insight that potential feelings of guilt can be useful as a means of self-control, especially if this potential is visible to others.

(ii) Esteem and disapproval. Esteem is a normative incentive that exists if a person cares intrinsically (in addition to instrumentally) what others believe about his behavior. Someone might gain utility directly from believing that others esteem him and lose utility from believing that others disapprove of him, regardless of whether these outsiders take actions that materially affect him. This effect of others’ beliefs on one’s utility is equivalent to saying that a person has a taste for others’ esteem. The idea is older than Adam Smith, but he put it well when he said in *The Theory of Moral Sentiments* that

> Nature, when she formed man for society, endowed him with an original desire to please, and an original aversion to offend his brethren. She taught him to feel pleasure in their favourable, and pain in their unfavourable regard. She rendered their approbation most flattering and most agreeable to him for its own sake; and their disapprobation most mortifying and most offensive. (Smith, 1790: 116).

Unlike utility from pride or guilt, utility from esteem or disapproval arises only when one believes other people have formed beliefs about one’s behavior. Disapproval can therefore be avoided by misbehaving secretly. As with guilt, the benefits of acting contrary to what others approve may outweigh the expected disapproval, especially when the disapproval is contingent on the offending behavior being detected. On the other hand, proper behavior is no guarantee of esteem, because esteem depends on one’s perception of other people’s beliefs, not on one’s own behavior. Not just good conduct, but other’s knowing about it – and knowing that they know about it – is necessary for esteem. And one may gain esteem without good behavior by fooling others into thinking one has behaved well.

Esteem and disapproval differ too from praise and censure, which are merely the expression of esteem and disapproval. Esteem and approval are subjective, based on beliefs about others’ opinions rather than on the actual opinions or their public declaration. Praise and censure are evidence of what others believe – but expression is
not necessary for an individual to believe that others have formed judgments of esteem or disapproval. The fact that actual expression is not required reduces the transaction costs of esteem as an incentive – though it also can lead to misincentives because of misperceptions. See Kuran (1995). Note, too, that praise and censure might also be valued for their own sake; one may value the expression even if it is already common knowledge that the speaker holds the expressed view, or even if it is common knowledge that the speaker is being hypocritical. The sweetest congratulation might be from a disappointed rival.


(iii) Shame. There is a third possibility. Some scholars, e.g. R. Posner & Rasmusen (1997), distinguish shame from guilt. Often, shame is used to mean the what we have termed disapproval, though with an emphasis on particularly intense and widespread disapproval. Shame might, however, mean something else: a negative emotion that arises from believing one has failed to meet standards set by the normative beliefs of others. On this account, shame falls between guilt and disapproval. Like guilt, it is an internalized sanction that occurs even if no one observes a norm being broken. Unlike guilt, the person feeling shame has failed to live up to the normative beliefs of others, which may be the case even if he has lived up to his own principles. As with disapproval, the standards of behavior are external, but unlike disapproval, the shamed person suffers disutility regardless of what others think. Suppose someone privately engages in sexual behavior X without feeling guilt (because he has not violated his own moral principles) or disesteem (because nobody else knows he has done it). Later he discovers that a friend strongly disapproves of X. The loss of utility that occurs only at this discovery would clearly not be guilt – by his own principles, he has done nothing wrong – nor disapproval – since the friend does not know that he has done X. We need a new category: shame. Likewise, there is a positive incentive analogous to pride or esteem if someone gains utility from successfully living up to the standards set by the normative beliefs of others, regardless of where whether he holds those same normative beliefs or whether others know he has succeeded (Cf. McAdams 1997:382-86). Shame and guilt are, of course subjects long studied in psychology. For entry into the literature, see Cosmides, Tooby & Barkow (1992), Harder (1995), and Tangney (1995).

b. Conventions

As noted above, many behavioral regularities that seem normative may in part or
whole be motivated by non-moral concerns, even when not driven by common tastes or fear of government penalties. These are conventions. Scholars in law and economics were analyzing social behavior driven by what we call conventions well before the word “norms” became popular, e.g., Brinig (1990) on wedding rings and Schwartz, Baxter & Ryan (1984) on dueling. A number of simple ideas from game theory can explain seemingly normatized behavior as driven by the usual incentives studied by economist, with no need to appeal to tastes.

One of the most important settings for conventions is the coordination game, in which the payoffs of the players are highest if they coordinate with each other. The problem is not conflicting desires but the need to avoid a discoordination that hurts everyone. This game leads the establishment of standards, whose importance is explained in Kindleberger (1983). A simple example is driving on the right side of the road.

Conventions also are important in repeated games, in particular when reputations can arise. Klein & Leffler’s seminal 1981 article on reputation essentially models it as an equilibrium of a repeated game in which a player is willing to forgo present profits in exchange for a good reputation that will yield him future profits. It may look as if a seller is providing high quality out of pride of workmanship or fear of disapproval, but he is actually motivated purely by material gain. Hirshleifer & Rasmusen (1989) use the idea of repeated games to explain ostracism – the expulsion of rule-breakers from groups, and Axelrod & Hamilton (1981) show the power of reciprocal altruism in “tit-for-tat.”

Signalling equilibria create still another form of convention. Someone may take a costly action to signal his inclinations or ability. This occurs if someone with baser inclinations or lower abilities would not be willing to bear the cost of the signal, whether it be the provision of advertising or restraint in taking advantage of the uninformed, a requirement known as the “single-crossing policy” because it can be formalized as requiring that the indifference curves in money-signal space of different types of agents cross only once (see Rasmusen, 2001, Chapter 12). For example, E. Posner (2000b; 2002) has explained a wide variety of behaviors as signals of one’s discount rate, which is important to revealing one’s suitability as a partner in repeated games (though see McAdams (2001) for a critique), and Fremling & Posner (1999) apply signalling models to sexual harassment law. Often, however, it is hard to tell which convention is at work – signalling information or reciprocating in a repeated game – as Kahan (2002) observes.

Sometimes conventions are formalized in the shape of institutions, as demonstrated by Ostrom (1990, 1991) in general, Cooter (1991) in the land system in New Guinea, and Milhaupt & West (2000) in organized crime. Institutions are rule-setting bodies that unlike government lack the power to coerce through the use of legal force but that can use conventions – involving ostracism, reputation, or information
transmission – to enforce their rules.

Since these convention models so often obviate the need to use norms to explain behavior, we will lay them out in slightly greater detail before proceeding to analysis of norms proper.

Coordination Games. In a coordination game, two or more players makes choices that will help them both if they match. Two drivers, Row and Column, may each need to decide whether to drive on the right side of the road or the left as they approach each other. The most important thing for each is that they make the same choice (which will mean that they avoid hitting each other). Assume it is also better if both choose to drive on the right, since they are driving cars with steering wheels on the left side. Table 1 shows the payoffs.

**Table 1: Ranked Coordination**

<table>
<thead>
<tr>
<th></th>
<th>Drive on Right</th>
<th>Drive on Left</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Row</strong></td>
<td>7, 7</td>
<td>0, 0</td>
</tr>
<tr>
<td><strong>Column</strong></td>
<td>0, 0</td>
<td>6, 6</td>
</tr>
</tbody>
</table>

Payoffs to: (Row, Column)

This game has two Nash equilibria if the choices are made simultaneously – (Right, Right) and (Left, Left). These equilibria can be Pareto ranked, but each is an equilibria. If each expects the other to drive on the Left, that is a set of self-fulfilling expectations in a simultaneous-move game. If the game were sequential, the only equilibrium would be for Row to choose Right and for Column to follow a strategy of imitating Row.

Many behavioral regularities are coordination games. Such behavioral regularities are often called norms, but not in our terminology because they are driven by simple self-interest rather than normative beliefs. Normative rules are not necessary to persuade people to avoid self-destruction in car crashes.

The Repeated Prisoner’s Dilemma. A second major category of convention model is the repeated prisoner’s dilemma. Unlike coordination games, prisoner’s dilemmas have complete conflict between the objectives of the players. In the classic story, two prisoners, Row and Column, are being questioned separately. If both confess, each is sentenced to eight years in prison. If both deny their involvement, each is sentenced to
one year. If just one confesses, he is released but the other prisoner is sentenced to ten years, as shown in Table 2.

Table 2: The Prisoner's Dilemma

<table>
<thead>
<tr>
<th></th>
<th>Deny</th>
<th>Confess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deny</td>
<td>-1,-1</td>
<td>-10, 0</td>
</tr>
<tr>
<td>Confess</td>
<td>0,-10</td>
<td>-8,-8</td>
</tr>
</tbody>
</table>

Payoffs to: (Row, Column)

The equilibrium of Table 2's game is (Confess, Confess), with equilibrium payoffs of (-8,-8), worse for both players than (-1,-1). Sixteen, in fact, is the greatest possible combined total of years in prison.

So far, no useful convention has emerged. But what if the game is repeated? Would the players arrive at a convention of choosing Deny in the early repetitions, knowing that they will be in the same situation in the future, with the possibility of revenge? Not if this is all there is to the game. Using an argument known as the Chainstore Paradox after its application to store pricing (where the Deny/Confess actions become Price-High/Undercut-Price), Selten (1965) explains that in the last repetition, the players will choose Deny because future revenge will be impossible, so in the second-to-last repetition the players will not have any hope for future cooperation, so in the third-to-last they will have no hope, and so on to the first repetition.

If the game is infinitely repeated, the Chainstore Paradox does not apply, and there exists an equilibrium in which the players choose Deny each time. Real-world interactions do not last forever, but Kreps, Milgrom, Roberts & Wilson (1982) show that with incomplete information, the addition of a small possibility of emotional behavior by a player such that he will choose Deny until the other player chooses Confess, can make (Deny, Deny) an equilibrium until near the last repetition. This is true even if the game does have a definite end, because if the other player does not know whether his opponent is emotional in this way or not, his best strategy turns out to be to treat him gently until late in the game. The infinitely repeated game with complete information is often used as a simpler model that comes to conclusions similar to those of the more realistic but more complicated finitely repeated game with incomplete information.

Signalling. The last type of convention model that we will describe here is the signalling game – one which is especially prominent in the norms literature because it is a central
idea in Eric Posner’s work, including in his 2000 *Law and Social Norms*. We will use a particular example from Rasmusen & R. Posner (1999), a model of employers preferring married to single workers. Suppose that 90 percent of workers are "steady," with productivity \( p = x \), and 10 percent are "wild," with productivity \( p = x - y \). Each worker decides whether to marry or not. Marriage creates utility \( u = m \) for a steady worker and utility \( u = -z \) for a wild worker. Employers, observing whether workers are married but not whether they are wild, offer wages \( w_m \) or \( w_u \) in competition with other employers, depending on whether a worker is married or not. We observe that \( w_m > w_u \).

We do not need norms to explain the higher wage for married workers. Employers have incentive to use marital status as a signal of productivity and to discriminate against single workers even if nobody thinks that marriage per se makes someone better or worse. The employer has no intrinsic reason to care whether the worker is married or not, since wild workers are less productive whether they are married or not. The only significance of marriage for the employer is its informational value as a signal of steadiness.

Unlike many signalling models, here there is only a single equilibrium. If \( z \) is large enough (greater than \( y \)), the employer will pay wages of \( w_u = x - y \) and \( w_m = x \). The steady worker will get married, and the wild worker will stay single. Steady workers will marry regardless of the effect on their wage, and wild workers will stay single even though they know that if they married an employer could be fooled into believing them to be steady – an example of the “single-crossing property” mentioned above.

The employers in this example might be unthinkingly obeying a rule of thumb of paying married workers more. Businessmen, like private individuals, follow many behavior rules without inquiring into their rationality. Following the rule is efficient and profit-maximizing even if no businessman understands its origin or rationale. When asked, an employer might say he pays married workers more because they deserve the higher wage, or need the higher wage, even though that is not the true reason. Thus, the convention of signalling is easily confused with a norm.

Signalling has implications for how laws should be designed. In this model, subsidizing marriage not only would be useless for raising productivity, but would lower it by depriving employers of useful information about the marginal product of their workers. Similar loss of information would occur if government forbade employers to use an applicant's marital status in making a hiring decision. Thus, in this model, it would be wrong for the government to start with the true premise that married workers are more productive and arrive at the conclusion that if more workers were married, productivity would rise; but it would also be wrong for the government to start with the equally true premise that a worker’s getting married has no effect on his productivity and arrive at the conclusion that it would make no economic difference if firms were forbidden to
discriminate by marital status.

Signalling models must be treated with care. They are “all-purpose” models that can “explain” practically any pattern of observed behavior give the right assumptions. The model above, for example, could as easily have been made a model in which steady workers derive less direct utility from marriage, in which case singleness would be the signal of ability, not marriage. This flexibility is both a strength and a weakness of signalling models.

Bayesian Learning in Cascade and Bandit Models. What seems to be norm-based behavior can also be entirely non-strategic, so neither norms nor conventions are needed to explain group behavior. One example is Rasmusen (1996), which explains stigma against the employment of criminals as arising from employer calculations of average ability based on population averages that can “tip” the level of criminality even if no single worker or employer thinks his own behavior will affect which equilibrium is played out. Another is the single decisionmaker “Two-Armed Bandit” model of Rothschild (1974), which shows how seemingly irrational, mistaken behavior can arise as the result of a rational policy of first investigating various possible behavior rules and then settling down to what seems best and never again experimenting.

A model of this type which has attracted considerable attention is the theory of cascades, originating with Banerjee (1992) and Bikhchandani, Hirshleifer & Welch (1992) and summarized in Hirshleifer (1995). It shows how fashions and fads may be explained as simple Bayesian updating under incomplete information, without any strategic behavior. Consider a simplified version of the first example of a cascade in Bikhchandani, Hirshleifer & Welch (1993). A sequence of people must decide whether to Adopt at cost .5 or Reject a project worth either 0 or 1 with equal prior probabilities, having observed the decisions of people ahead of them in the sequence plus a private signal. Each person’s private signal is independent. A person’s signal takes the value High with probability $p > .5$ if the project’s value is 1 and with probability $(1-p)$ if the project’s value is 0, and otherwise takes the value Low.

The first person will simply follow his signal, choosing Adopt if the signal is High and Reject if it is Low. The second person uses the information of the first person’s decision plus his own signal. One Nash equilibrium is for the second person to always imitate the first person. It is easy to see that he should imitate the first person if the first person chose Adopt and the second signal is High. What if the first person chose Adopt and the second signal is Low? Then the second person can deduce that the first signal was High, and choosing on the basis of a prior of .5 and two contradictory signals of equal accuracy, he is indifferent – and so will not deviate from an equilibrium in which his assigned strategy is to imitate the first person when indifferent. The third person, having
seen the first two choose *Adopt*, will also deduce that the first person’s signal was *High*. He will ignore the second person’s decision, knowing that in equilibrium that person just imitates, but he, too will imitate. Thus, the first person’s decision has started a cascade, and even if the sequence of signals is (*High, Low, Low, Low, Low...*), everyone will choose *Adopt*. A “cascade” has begun, in which players later in the sequence – starting with the second one in this example! – ignore their own information and rely on previous players completely. We have chosen an extreme example, in which the cascade starts immediately with probability one, but the intuition is robust, and more complicated models yield interesting implications for how signal quality and correlation affect the probability of a cascade starting.

Learning models such as these are useful for modeling apparently irrational behavioral regularities. Suppose we observe a culture that tries to cure malaria by bleeding the patient. This does not have to be the result of norms. Rather, it may be that after trying other methods and failing– perhaps even trying quinine bark without consistent success – the tribe has rationally if mistakenly settled down to bleeding as the best method based upon available evidence. But this is neither a norm in our sense nor a convention, since it is the result of neither obligations nor strategic interactions.

Conventions interact with normatized incentives. Kreps, Milgrom, Roberts & Wilson (1982) show how just a few people with normatized incentives can lead many others to imitate them in their behavior. Kuran’s 1995 *Private Truths, Public Lies: The Social Consequences of Preference Falsification* shows how such deception about one’s true preferences can lead to sudden reversals of public opinion. One of Kuran’s examples is the collapse of communist regimes when most citizens suddenly discovered that the support for the regime was not genuine but based on a complex “web of lies.” Similarly, Kuran & Sunstein (1999) propose that informational and reputational cascades sometimes combine to cause stampedes toward ill considered regulations, and Kuran (1998) analyzes the interaction between cascades and norms in the context of ethnic identification. Kuran and Sunstein distinguish “informational” cascades of the sort just described and “reputational” cascades that occur only because individuals expect to gain by being known to conform. In our terminology, informational conformity is one way to produce a convention; reputational conformity is one way to produce a norm.

Thus, after discussing such diverse convention models as signalling, repeated prisoner’s dilemmas, and cascades, we see that much of human behavior that seems to be driven by moral beliefs is actually driven by utility maximization in the narrow sense of Holmes’s bad man, though by a bad man sophisticated enough to know how important strategic behavior is to his success. La Rochefoucauld said, “Hypocrisy is the homage vice pays to virtue.” In the present context, “Convention is the homage *homo economicus* pays to norms.”
c. The Origin of Norms

Although we have shown how a variety of apparent norms could actually be conventions, the study of conventions is important to norms more than just for explaining them away. While we have distinguished between conventions that work by appealing to standard, non-normative tastes and norms that work only when supported by feelings of obligation, there remains the question of where those feelings come from. Conventions are an important part of the answer.

How people come to have any of the three normative drives just discussed – guilt, esteem, or shame – is a subject given considerable attention by biologists ever since Darwin’s 1874 *Descent of Man*. For the biologist, any kind of tastes, standard, or norm is the result of an equilibrium, an evolved outcome of a process similar to maximization, although less calculated and with results harder to call “optimal” in a meaningful way. Biologists have also studied what would be conventions in humans (because motivated by calculation) but are commonly norms in animals (because motivated by preferences – the inborn preferences we call instinct). Though genes are “selfish,” E. Wilson (1980) shows that there are conditions under which helping behavior is necessary to survive – e.g., hunting large prey, warding off predators, etc. – and motives such as guilt, shame, or esteem may induce such helping behavior. See Trivers (1971); Jack Hirshleifer (1978, 1987); Fershtman and Weiss (1998). The approach has been picked up in ethics (Peter Singer’s 1980 *The Expanding Circle: Ethics and Sociobiology*), anthropology (Boyd & Richerson’s 1988 *Culture and the Evolutionary Process*), political science (Ostrom (1991) and her 1990 *Governing the Commons*), and economics (Jack Hirshleifer 1978), Bergstrom (2002) and Sartorius (2002) generally; Cameron (2001) on sexual behavior).

The biological approach is really an extension of the idea that humans are born with certain norms instilled in them (e.g. Romans 1, Aristotle’s *Nichomachean Ethics*, Aquinas’s *Summa Theologica*), an idea that under the name of “natural law” is the subject of a quite different branch of scholarship (e.g., James Q. Wilson’s *The Moral Sense* (1993); Budziszewski’s *Written on the Heart: The Case for Natural Law* (1997); the essays in George (1995)). The biological approach, however, with its analytic framework of evolution as a source of utility functions, has proven more useful than natural law as a source of explanations in law and economics.

Biological evolution brings to mind the literature in law and economics on the evolution of the common law, for which Rubin (1977) and Priest (1977) provide seminal
articles and Zywicki (2003) summarizes and gives historical detail. The common law is a special example of customary law, and in primitive societies, even more than modern ones, it can be difficult to distinguish between norms and laws – between what is enforced by guilty, esteem, and shame, and what is enforced by the power of the state. In medieval Europe the function of the government was not to make law, but to discover it, as Hayek discusses in Volume 1, Chapter 4 of his 1973 *Law, Legislation and Liberty*. It was natural for Hayek to precede his discussion of laws in that work with discussions of the evolution of norms. What his verbal discussion means and whether it is correct is controversial, as detailed in Whitman (1998), but the basic project is a sound one: to examine how norms evolve and the extent to which group selection favors desirable norms.

We should emphasize, however, that norms need not always be preceded by conventions. For example, Pettit (1990) and McAdams (1997) claim that a new pattern of approval and disapproval can create a new behavioral regularity, given a desire for esteem. A norm arises when individuals desire esteem and these three conditions hold: there is a strong pattern of approval or disapproval for a given activity, there is a risk that others will detect one’s engaging in the activity, and there is something approaching common knowledge of the approval pattern and risk of detection. Geisinger (2002) and McAdams (2000) claim that law can facilitate the process of norm emergence by publicizing the existence of a new consensus.

4. The Importance of Norms to Legal Analysis

In this section we describe how the existence and operation of norms affect the positive and normative economic analysis of law and legal institutions.

a. Positive Analysis: How Norms Affect Behavior

(i) Generally. Norms matter to the positive economic analysis of law in two respects: in predicting how a change in legal rules affects behavior, and in explaining how law is made.

One cannot accurately predict behavior without knowing something about all the incentives that influence behavior – which includes normative incentives – as well as the way that legal change interacts with them. Economic analysis of law needs to consider carefully how norms may govern behavior in the absence of law and how a new legal rule may intentionally or unintentionally change (or fail to change) a norm.

Norms are, of course, highly diverse – as diverse in application as laws. Ellickson
(1991: 132) has usefully categorized rules of any kind, including norms, into five groups. *Substantive* norms concern the conduct that is to be regulated in the first place, and the other four categories are ancillary. *Remedial* norms prescribe penalties or rewards for norm violation, *procedural* norms determine how information about violation is to be collected and used, *constitutive* norms govern how norms are created, and *controller-selecting* norms divide the labor of social control among different people.

Consider the norm of property in snowy-weather parking spots in Chicago described by Epstein (1999). The substantive norm says that only the person who dug the snow out of a spot is entitled to park there, whereas others who park will suffer guilt, disesteem, shame, or more concrete sanctions (Mayor Daley said, "I tell people, if someone spends all that time digging their car out, do not drive in that spot. This is Chicago. Fair warning."). Epstein does not describe the ancillary norms, but let us imagine what they might be. The remedial norm might be that someone who parks in the wrong spot will have his car window broken. The procedural norm might require that the enforcer make some attempt to find and warn the violator before he resorts to violence. The constitutive norm might be that the norm can be changed only by explicit agreement of the residents of a street, and the controller-selecting norm might be that only the “owner” of a space is allowed to punish the violator.

Ignoring norms (or conventions) can cause one to overstate the significance of law, as suggested by the comments of Mayor Daley, the official ultimately in charge of enforcing both parking and vandalism laws for the City of Chicago. Norms matter in several ways. First, economists sometimes assume that a legal rule influences behavior, when an empirical investigation would show that the legal rule has no influence because group norms exclusively govern the behavior. Ellickson (1986, 1991) famously found that ranchers in Shasta County, California ignored legal rules concerning animal trespass and resolved disputes over cattle trespass damages according to “neighborly” norms, even though they had the legal right to go to court. Indeed, often one group norm is that member should never make use of their legal rights. For similar results concerning workplace norms (or conventions) and law, see Kim (1999) and Rock & Wachter (2002). Either can be strong enough to trump laws. Second, one might think a legal rule a necessary condition of some observed behavioral regularity when a norm would maintain the same (or nearly the same) regularity without the law. For example, a norm that promises must be kept might, in identifiable circumstances, produce as much promise-keeping as legal liability, or at least enough so as to make the costs of legal enforcement no longer worthwhile (Macaulay 1963; Scott 2003). Third, one might overestimate the ability of legal change to produce a behavioral change by underestimating the degree to which the existing behavior is driven by norms (Kahan 2000).

On the other hand, ignoring norms can also cause one to understate the
significance of law. Economists sometimes assume that a legal rule is not necessary to change behavior when on closer analysis they would find that without new laws, norms will freeze the behavior in place. For example, market competition might not eliminate race discrimination if social norms require such discrimination (McAdams (1995)). Moreover, changing a law might have a greater effect if legal sanctions work not just directly, by raising the price of a behavior, but indirectly, by changing norms. A new law might change perceptions of what incurs disapproval (McAdams 2000), create a new basis for shame, or even change a person’s own preferences and create guilt as Dau-Schmidt (1990) discusses in the context of criminal law. Kahan (2003) writes of the pervasive norm of “reciprocity,” which he believes underlies much mutually productive cooperation in both small groups and society generally, but notes many ways that law can unintentionally undermine or intentionally facilitate such reciprocity. The extent to which the law actually does affect norms – and the ease with which such claims for new laws can be made – is an interesting question discussed in number of articles, e.g., Posner & Rasmusen (1999), Picker (1997), Hetcher (1999), Dharmapala & McAdams (2003); McAdams (2000); and Kahan (2000). Ellickson (2001) addresses the issue by comparing the ability of government and private “norm entrepreneurs” to change norms. Empirical work is harder to come by, but see Massell (1968) on law and change in Soviet Central Asia.

Positive analysis of law also seeks to explain a second point: why particular law-making institutions – the legislature, courts, or administrative agencies – create particular laws. Often one cannot fully explain the existence of a law without understanding the norms that give rise to it, or the absence of norms that would block it. Where public choice theory emphasizes the material interests citizens have in enacting or defeating legislation, attention to norms reveals that many people are highly motivated to create rules that do not affect their material interests. A person who believes that certain behaviors are immoral – e.g., pornography, abortion, flag burning, animal testing, or environmental exploitation – often favors laws forbidding or restricting such behavior. In turn, in a democratic system, such people’s votes give the legislature incentive to enact laws supporting the norm. Or, if the political system gives him enough slack, the legislator, judge, or administrator may use his power to enforce the behavior he views as morally required.

Why would voters or lawmakers believe a law is necessary if the behavior is already enforced by guilt, disapproval or other normative incentives? An obvious reason is that the existence of a norm does not imply perfect compliance. Many people will occasionally face situations where the expected benefits of norm violation exceed the expected costs, and certain people may never obey the norm because they feel no guilt from violating it and can avoid detection. Another reason, more in keeping with public choice theory, is that even the norm, much less compliance, might not be universal.
Different lawmakers will push to enforce the norms of the groups that support them, norms which come into conflict just as much as budget priorities, and often with more bitterness because of the normatized preferences of each group and the difficulty of compromise. It is hard to “split the difference” on abortion.

Still another motivation for laws as a supplement for norms is that the lawmaker may gain from purely symbolic endorsement of a norm, even if that endorsement is not expected to change behavior. There may be no observed flag-burning in a jurisdiction with strong patriotic norms, but voters may want to go further and express their disapproval by a symbolic declaration. Indeed, it is all the easier to pass such a law if nobody in the jurisdiction actually does want to burn a flag so no resources would have to be devoted to enforcement. Closely related is the function of laws as helping to create and perpetuate norms – one of the “expressive” functions of law discussed in Dharmapala & McAdams (2003), Geisinger (2002) and McAdams (2000). By saying what people should do, even if there is no penalty, the law tries to shift or maintain tastes, and to educate a society’s newcomers – children and immigrants – in its norms. Law may serve the same function as the “rituals” that Cappel (2003) discusses, reinforcing attitudes by aiding communication of what is esteemed or by actually changing tastes by changing habits (on which see the experiments in Wells & Petty (1980): subjects instructed to nod their heads “Yes” repeatedly while listening to someone speak came to agree more with what the speaker said).

(ii) Specific Norms Regarding Law. Besides these general points, some specific law-related norms have particular relevance for positive legal analysis. The most important are the norms of “legal obedience” – that people should obey the law – and “the rule of law” – that laws should be knowable in advance rather than the purely discretionary decision of some authority.

People often feel obliged to obey laws, or at least laws they perceive to be “legitimate” from the very fact that they are laws, rather than from any other motivation. These people suffer guilt, shame, or disapproval from breaking the law. The norm of legal obedience provides an incentive to obey the law that is independent of material sanctions (though if it is based on esteem and disapproval it still depends on violations being detected). This effect is particularly important for offenses that are malum in prohibitum – wrong only because illegal – because the prohibited act is not itself governed by a norm and the only relevant norm is legal obedience. One should not bring more than $10,000 in currency into the United States without declaring it on the customs form, but only because it is a legal wrong. By contrast, the norm against malum in se offenses such as murder is independent of its illegality. One should not kill unjustifiably, because that is a moral wrong – which also happens to be a legal wrong. (Other acts may be malum in se but not malum in prohibitum, e.g., adultery in the United States of 2004, as discussed in
Related to the norm of legal obedience is the ideal of “the rule of law.” Defining this norm is difficult, but a central element is the idea that, as Fallon (1997:3) puts it “the law – and its meaning – must be fixed and publicly known in advance of application, so that those applying the law, as much as those to whom it is applied, can be bound by it.” This norm constrains government officials who wield official power, a non-legal sanction against their illegal use of discretion or violation of rules. Thus, the rule of law is contrasted with “the rule of men.” The norm of the rule of law is of great significance to how well laws work, since the alternative is costly, perhaps prohibitively costly, monitoring of executive and judicial officials. Development economics is by now quite conscious that it is not enough to establish good laws; one must, in addition, eliminate corruption and enforce laws fairly. See Rose-Ackerman (1999); Brooks (2003). We will not go into the large topic of jurisprudence, but merely note that it is an area in which the law-and-economics of norms might be usefully applied.

Other norms govern specific legal actors. To understand how a legal institution works, one must understand the norms governing that institution. For example, given how central the jury is to the legal system, it is odd how little attention economists have paid to the fact that jurors are paid by the day (and frequently less than their forgone wage) rather than based on the quality of their understanding or resolution of the case. Pettit (1990) notes that without normative motivations the successes of juries are puzzling, but that with such motivations we may explain why jurors pay attention to evidence, deliberate, and vote according to their evaluation of the evidence.

Similarly, other group norms besides the norm of the rule of law appear to be important – if not entirely effective – in constraining the behavior of judges, legislators, prosecutors, police, and other executive branch officials. A particularly interesting set of law-related norms are those governing lawyers. Many of the ethical rules governing lawyers lack genuine sanctions and may be understood as efforts to strengthen or create professional norms (Painter 2001; Wendel 2001 – though see Fischel 1998 for a more skeptical view). There is some empirical evidence that norms even constrain the fees lawyers seek (Baker 2001), though here there may be difficulty separating norms from convention or private rules. An interesting example is *Goldfarb v. Virginia State Bar*, 421 U.S. 773 (1975), which concerns whether a bar association may expel members who charge a fee below a posted minimum as unethical. *Goldfarb* examines an industry group that seeks to regulate itself with professional norms, which shade into organizational rules and then into law.

(iii) Specific Laws Regarding Norms. Often one cannot understand the meaning of a specific legal rule without understanding the norms or conventions to which the rule
explicitly or implicitly refers. At one extreme, law simply incorporates certain customs *in toto*. Cooter & Fikentscher (1998:315) gives an example:

> When making Indian common law, tribal judges confront a central problem in legal anthropology: How to distinguish customary obligations that are enforceable at law (which can be called "common law") from customary obligations that are not enforceable at law (which can be called "mere customs")? Put succinctly, the problem is to distinguish "law from custom." If a custom is law, then legal officials are obligated to enforce it, whereas if custom is not law, then legal officials require an independent justification for enforcing it.

Obviously, in this case, one cannot know the content of the law without knowing the content of the custom (convention or norm) it enforces.

Norms are also used more narrowly, to flesh out statutes or judge-made law rather than to create laws out of whole cloth. Rather than fully specifying a substantive standard, many legal rules and doctrines “incorporate by reference” existing customs or practices, which in some contexts means norms and in other contexts means conventions. Legal definitions of obscenity explicitly incorporate local “community standards” (see *Jenkins v. Georgia* 418 US 153 [1974]). Given the strong normative attitudes about the depiction of sex acts, the “standards” that the law incorporates are norms. Also, various torts – e.g., battery, invasion of privacy, and intentional infliction of emotional distress – include open-ended elements such as outrageousness or the absence of a customary privilege that implicitly incorporate norms. Both the crime and the tort of battery refer in part to the “offensive” touching of a person, which refers to norms. Other rules incorporate norms only indirectly or implicitly. Defamation law determines that certain statements are defamatory per se because they presumptively hurt the individual’s reputation. What is defamatory per se is often the accusation of a norm violation – *e.g.*, accusing a person of committing adultery. In recent years, changing attitudes towards homosexuality have made norms a subject of interest to courts trying to determine what is defamatory – see, for example, *Donovan v. Fiumara*, 114 N.C. App. 524 (1994). In many statutes, the crimes of extortion, coercion, or blackmail includes the threat to reveal any secret that will tend to expose the victim to “hatred, contempt or ridicule,” which often includes the threat to reveal a norm violation. See Model Penal Code § 223.4 (1985) (Theft by Extortion) and § 212.5 (Criminal Coercion). A full understanding of the content of the law in these cases (and others) must correctly understand the content of a norm. A positive analysis of the consequences of the legal rule must also consider possible dynamic effects of incorporating the norm into the rule.

How does the normative analysis of law need to account for norms? Broadly speaking, there are two issues. First, how should welfare analysis incorporate the existence of norms and normative incentives? Second, when are norms efficient – or, more to the point, when are they preferable to law as a way to regulate behavior?

(i) Welfare analysis. Norms change the welfare calculus in several ways. First, we must incorporate guilt, esteem, shame and pride into welfare via their direct effects on utility. Kaplow & Shavell 2001a, 2002b examine how the normative incentives of guilt and pride (which they term “virtue”) affect the welfare analysis of legal and moral rules. Ideally, there is a set of guilt and pride inclinations that ensures optimal behavior by each individual. Individuals sufficiently motivated would act optimally and therefore never have to incur guilt, which otherwise decreases welfare. But Kaplow & Shavell introduce reasonable constraints that complicate the analysis: that the process of inculcating guilt and pride is costly, that there is some psychological limit to the degree of guilt or pride individuals can feel, and that guilt or pride can be inculcated only for broad “natural” groupings of acts – such as lying – rather than for each particular act depending on its welfare effect – such as an inefficient lie. The result is a series of interesting tradeoffs between the use of guilt and pride and the optimal groupings of acts. Shavell (2002) then examines the optimal tradeoff between the use of these moral motivations and the legal system. The advantage of morality is that, compared to law, it is cheap and its internal incentives work without the external detection of anti-social acts. But the legal system can impose rules involving finer gradations in conduct than guilt-enforced morality, can change the rules more quickly than morality in response to changed circumstances, and can usually impose higher sanctions for the most tempting suboptimal acts.

Second, people who can feel guilt or pride in their own behavior will likely feel similar emotions as a consequence of observing the behavior of others, including that of government agents acting on their behalf. Thus, individuals may believe that certain legal outcomes are “fair” or “unfair,” and thereby gain or lose utility from observing the outcomes. If so, the welfare analysis of legal rules must account for these effects on utility. For example, several theorists, such as Polinsky & Shavell (2000), Sunstein, Schkade & Kahneman (2000), and Kaplow & Shavell (2002), consider the significance of the popular view that punishment should be proportionate to the crime. Those holding this view may suffer disutility if maximal sanctions are imposed for non-serious offenses. Consequently, even if maximal sanctions (and minimal levels of detection) would otherwise be socially optimal, they might be suboptimal once we include the disutility of disproportionality. This argument is open to abuse, since it can be called in to defend any policy that some people favor, but that does not so much diminish its validity as call for empirical validation of claims that the utility effect is large enough to matter. Moreover, it provides one way to interpret the “retributive” function of punishment: observers feel utility when they observe misbehavior punished proportionately, and feel disutility when
misbehavior receives disproportionately low punishment, including the extreme case of escaping punishment altogether.

Third, normative analysis must address the question of whether the social objective function should incorporate norms merely via their effect on the utilities of individuals or in addition to those utilities. If the objective function maximizes utility, it will take into account, for example, the distress that people feel at what they consider to be “unfair” outcomes, but the social planner might wish to reduce “unfairness” even beyond the effect on utilities. This double-counting might be legitimate, but the analyst should be aware of what he is doing, and double-counting necessarily means abandonment of Pareto optimality, which is an important argument against it. See Kaplow & Shavell (2001, 2002a). The logic of conventional welfare economics, with its criteria of efficiency or wealth maximization, requires instead that norms enter via their effects on utilities. Richard Zerbe’s 2001 book, *Efficiency in Law and Economics*, is useful in clarifying this, and in showing how norms in utility functions can be operationalized by looking at a person’s willingness to pay to have a norm obeyed.

All three of these welfare considerations treat norms as exogenous. We previously noted, however, that the law might in some cases influence norms, which can involve (when guilt inclinations are changed) a change in tastes. The welfare analyst must therefore decide how to deal with the possibility of preference change. Economists since Strotz (1955) have studied the problem of how to do welfare analysis when tastes are variable or when people are poorly informed. For example, legal rules against race discrimination might initially generate direct utility for people who regard such rules as fair and disutility for people who regard them as unfair. But if the rules produce a change in preferences over time, diminishing the internalized norm of discrimination, static analysis will be misleading (Kaplow & Shavell (2002)). If the analyst knows that an anti-discrimination law will lead an individual to change his preferences and actually prefer the law after five years, should the individual’s present preferences trump his future preferences? Should this be the case even if the individual knows how he will change? Such questions have been much discussed in various contexts; see Dau-Schmidt (1990), Kuran (1995); and Ng (1999).

Cooter (1998) links norms to the concept of a “Pareto self-improvement”: an individual who perceives the advantage of having different preferences, even from the vantage point of his existing preferences, may work to change his preferences. For a critique, see McAdams (1998). If people are poorly informed, however, there can be a conflict between maximizing their utility ex ante – making the choices that they, with their poor information, would make – and ex post – the choices they would have made if well informed. Richard Posner applies this idea in his 1992 *Sex and Reason*, in which he suggests that norms against sexual practices such as homosexuality would disappear if
their holders had better information. The big practical problem, of course, is determining whose information is wrong, since each side may well believe that the other’s beliefs are sincere but misguided.

Welfare analysis of preference change is particularly complex in the case of interdependent utility functions. The norm of retribution, for example, may be supported by preferences in which one derives utility from the disutility of another – the offender (see Kahan (1998)). By contrast, altruism may underlie norm of gift-giving, which, as Kaplow (1995) explains, increases the utility both of the individual holding the norm and of others on whom he acts. John Stuart Mill is hostile to what he calls “other-regarding” preferences in On Liberty (1859), though as has been pointed out in Kaplow & Shavell (2002) and James Fitzjames Stephen’s Liberty Equality Fraternity (1873), his own tone is highly moralistic, and sufficiently obscure that it is hard to make sense of how he decides which nonmaterialistic preferences are legitimate and which are not.

(ii) Norms versus Law. The second broad issue is whether norms, or certain identifiable classes of norms, are generally efficient or inefficient. This matters to whether the coverage of law should be expanded or shrunk.

Norms have the obvious advantage of low transactions costs compared to law. They do not require police, courts, collection agencies, or prisons. If they are fully internal, they do not even require detection. Thus, they seem particularly appropriate for regulating externalities too small to justify appeal to the courts, or for those whose detection and proof are particularly difficult. On the other hand, norms are trickier to create than laws, and are not typically the subject of policy discussion. Rather, the usual question is whether society should create laws to supplant norms.

Legal authorities will often wish to defer generally (rather than case by case) to norms in domains where norms are more efficient regulators of human conduct than legal rules. Thus, one needs to compare the efficiency of legal rules to decentralized norms. Eric Posner (1996) examines the case for deferring to norms in groups governed by them. Legal regulations intended to protect individuals may have the unintended consequence of lowering the value of group membership, thus weakening the power of groups to enforce their norms. Thus, the efficient legal rule might be one of non-interference. Shavell (2002) makes a general comparison of the comparative advantages of law and morality, where morality includes both internalized and non-internalized norms.

Whether norms are generally efficient, or even efficient in identifiable circumstances, is contested. Everyone acknowledge the existence of dysfunctional norms, but Ellickson and Cooter, to take two major figures in the literature, are optimistic about the efficiency of group norms that affect only the members of the group, viewing norms
as mechanisms for deterring behavior with negative externalities and encouraging behavior with positive externalities. Both are pessimistic about norms between groups, where there is no incentive to account for the external effects. See also McAdams (1995) on the stability and inefficiency of norms of racial discrimination.

Others are more pessimistic about norms, even when they apply only to the group in which they arise. Several theorists make the general point that law is often superior to norms or conventions. Kaplow & Shavell (2002) argue that norms of fairness usually enhance social welfare by curbing self-interested behavior with negative externalities, but nonetheless conclude that norms are in many particular cases inferior to regulation by the optimal legal rule. They emphasize two disadvantages of norms that arise because norms are frequently inculcated in children and supported by feelings of guilt. First, the norm is often simpler than the optimal rule (e.g., never break a promise, rather than never inefficiently break a promise). Second, the norm is hard to change when new conditions make a different rule optimal. Kahan (2000) also emphasizes the stickiness of obsolete norms. McAdams (1997) raises the possibility that groups will enforce “nosy” norms that regulate behavior with only mild externalities. Norms may demand conformity to the other-regarding tastes of the majority even when the minority loses much more by frustration of its self-regarding preferences than the majority gains (e.g., regarding mate-selection criteria).

Similar points apply to conventions. Eric Posner (1996b, 2000) identifies various problems arising from poor information or strategic behavior that can make conventions and norms inefficient, justifying a corrective or supplementary legal rule. Mahoney & Sanchirico (2001) use evolutionary game theory to explain how the fittest convention in a given environment often deviates from the efficient one. See also Horne (2001) and Kübler (2001). In the case of either norms or conventions, of course, we must keep in mind that just because norms are inefficient does not mean laws would be efficient, any more than market failure in standard economic markets means that government regulation would be optimal rather giving rise to government failure instead. The issue is which is the greater danger, the purposeless inefficiency of norms or the purposeful inefficiency of law.

A second possibility is that efficient norms are “fragile” and therefore require not just non-interference but affirmative legal protection. This might justify otherwise puzzling rules of market-inalienability. Why does the law constrain the sale of parental rights and child labor? One possibility is that parental norms are a more efficient regulator of parenting practices than is law, but that normative incentives are weak compared to market incentives and that parenting norms would unravel if parents were fully subject to market incentives. An unregulated market would, on this view, leave parents to make individually maximizing but socially inefficient decisions about their
children, such as the choice to curtail their education in order to exploit their short-run potential in the labor market. Rather than overcome this problem by directly regulating the precise boundaries of parental conduct, however, one might instead enact rules of market-inalienability, constraining the operation of market incentives on parents, and leaving them subject to relatively more powerful normative incentives. This idea is distinct from but related to the idea that monetary incentives might “crowd out” non-monetary incentives. (Frey 1994).

5. Specific Applications.

a. Tort law. Shavell & Kaplow (2002: 134-43) note that there is a strong norm to avoid injuring others and to compensate them for injuries one does cause. See also Smith (1790: 104). Consistent with their general thesis, they claim that these norms generally improve social welfare but that law can make additional gains. Tort laws and litigation processes have certain advantages over norms in collecting the relevant information, imposing more optimal rule complexity, changing rules more quickly in response to changed conditions, and imposing greater sanctions. Compensation norms, however, also encourage lawmakers to compensate via tort law, and may therefore hinder reliance on an insurance regime when it is a better means of compensating accident victims.

Custom has long played a key role in tort law because it helps to decide whether an injurer was negligent. Since negligence is closely allied to failure to fulfill an obligation, negligence is, in the terminology of this chapter, the violation of a norm. As Hetcher (1999) explains, two possible uses of custom are as a per se rule, under which custom is a complete defense (exemplified by the leading case of Titus v. Bradford, 20 A. 517 (Pa. 1890)), and as an “evidentiary rule,” under which custom is only evidence about what is non-negligent. Justice Holmes preferred this second use, and said in Texas & Pacific Railway Company v. Behymer, 189 U.S. 468, 470 (1903), “What usually is done may be evidence of what ought to be done, but what ought to be done is fixed by a standard of reasonable prudence, whether it usually is complied with or not.” Judge Learned Hand’s opinion in The T. J. Hooper, 60 F2d 737 (2d Cir. 1932), is similar, adding the idea that norms may become outdated because of technological advances—the invention of the radio in that case (on which see Epstein (1992b)). Hetcher, however, argues against the modern preference for the evidentiary rule using the idea of the coordination game.

Norms are also important to what damages are awarded in tort. Cooter & Porat (2001) address the question of whether legal damages ought to be adjusted for the
normative penalties that an injurer has paid for his misdeed. Cooter (1997) argues that norms are central to whether punitive damages are awarded, though not useful to deciding their magnitude.

**b. Contracts and commercial law.** Kaplow & Shavell (2002: 203-13) review some evidence for the existence of a strong norm of promise-keeping, supported by guilt. Macaulay (1963) first documented that businesses do not rely on exclusively or even primarily on law to enforce agreements. They use norms, which may most simply be reduced to two: “(1) Commitments are to be honored in almost all situations; one does not welsh on a deal; (2) One ought to produce a good product and stand behind it” (p. 63). Cooter & Landa (1984) find that ethnically homogenous minority groups often dominate certain “middleman” positions in the markets of nations in which judicial enforcement of contracts is weak or non-existent. They claim that ethnic ties create informal enforcement mechanisms (norms or conventions of promise-keeping) that substitute for state enforcement of contracts. See also Landa (1981, 1994) and Davis, Trebilcock & Heys (2001). In a series of meticulous studies, Bernstein (1992, 1996, 2001) finds that many merchants groups prefer to enforce contracts, when disputes arise, through private trade association mechanisms, rather than rely on the state; see also Richman (2004) on Jewish diamond merchants in New York.

In cases of informal enforcement, norms do nicely in handling clear cases, but problems do arise because it is hard for a norm to sharply define when one “honors a deal” or “produces a good product.” Kaplow & Shavell (2002) suggest that although promise-keeping norms improve social welfare by making certain trades uniquely possible, often the optimal legal rule would do even better. In particular, norms are often too simple and their penalties are too weak to deter misbehavior when the stakes become high.

Much of what appears to be norms in business may be driven by the convention of the infinitely repeated prisoner’s dilemma. A particular important variant of this convention is reputation, as laid out in the classic article of Klein & Leffler (1981). We will recast their idea here as the formal model used in Chapter 5 of Rasmusen (2001).

Suppose we have sellers who can produce either a high-quality good at cost $c$ or a low-quality good at cost 0. Buyers all value the high-quality good at some amount much greater than $c$, and the low-quality good at 0, but cannot tell quality until after they have bought the product. The players make choices each period of what quality to choose, what price, $p$, to charge, and whether to buy or not, with a small discount rate of $r$ between periods and no end period.

If there were only one repetition, the unique equilibrium would be for the sellers to produce low quality and for the buyers to decide not to buy. In the repeated game, low
quality will remain the equilibrium outcome if expectations are pessimistic, but if the buyers believe that a given seller is reputable and will produce high quality, that too can be a self-fulfilling expectation.

In the high-quality equilibrium, a seller is willing to produce high quality because it can then sell at a high price for many periods, but if it produces low quality, the one-time savings in production costs is offset by the loss of future returns. Thus, an essential part of the model is that the equilibrium price be well above marginal cost. For the seller to produce high quality, its one-time gain from cheating and producing low quality – the revenue of \( p^* \) – must be no greater than the present discounted value of the alternative long-term profit of \( (p^*-c) \) each year, a value equal to \( (p^*-c)/r \). This requires that \( p^* = (1+r)c \). Any seller selling at a price higher than \( p^* \) would be undercut by a seller that sold at \( p^* \), but any seller selling at a price less than \( p^* \) would get no customers, since buyers would realize that such a low price does not provide enough incentive to keep quality high. Buyers rightly do not trust a seller who is not charging a high enough price. The reputable sellers then make positive profits because undercutting each other drives away customers.

Thus, a convention – that quality be high and any deviant firm be punished by future boycott – results in high quality despite the lack of immediate observability or enforcement by laws. The model can be applied to many situations of good behavior seemingly enforced by norms. In particular, under the name of “efficiency wages” it can explain high wages and honest behavior of employees in industries where trust is important (see Akerlof & Yellen (1986)).

Custom has played an important role in contract law at least since the time of Lord Mansfield. The Uniform Commercial Code says in Section 1-201(3) that an agreement is to be interpreted “by implication from other circumstances including course of dealing or usage of trade or course of performance” and numerous other sections of the UCC, listed in Bernstein (1999: note 1), follow this “incorporation strategy.” Custom here is more important as a convention than as a norm, and often its role really is just as an aid in interpreting a contract term; the use of custom as evidence of meaning is uncontroversial. Customs may become normatized, however, in a relatively simple way: someone who violates a norm is considered to be cheating in the same way as someone who lies or breaks a law, and the victim of the violation feels a visceral response.

Bernstein (1996) makes a useful distinction between “relationship-preserving” norms, which apply to continuing good relationships between businesses, and “end-game” norms, which apply to distressed relationships that are winding down. She notes that courts should not fall into the mistake of using relationship-preserving norms as a guide to how businesses would wish their affairs to be conducted once the relationship is
ending, and uses arbitrations of the National Grain and Feed Association as an illustration. Those arbitrations are heavily formalistic, she shows, in contrast to the UCC approach, and do not consider elements such as “good faith” in making decisions.

Naturally, norms vary from industry to industry, making this fertile ground for empirical study. Drahozal (2000), for example, argues that international commercial arbitration is more like the U.C.C. than the arbitrations studied by Bernstein. A leading case in international commercial arbitration, *Pabalk Ticaret Limited Sirketi v. Norsolor S.A.*, ICC Award of Oct. 26, 1979, No. 3131, 9 Y.B. Commercial Arb. 109 (1984), applied “the international lex mercatoria,” which the court said included a principle of “good faith which must preside over the formation and performance of contracts,” so that a party would be liable because of its faithless conduct. A requirement of “good faith” sounds like and may reflect moral obligations – norms – rather than mere convention. This example illustrates a problem running through the literatures both of norms and conventions: these rules are no more likely than laws to be universal, so case studies, requiring considerable study to yield single scholarly papers, are often more useful than either statistical or theoretical articles.

c. Corporate law. Corporate law has been the object of a surprising amount of scholarship on norms, considering that the corporation itself has no shame, guilt, or appreciation of esteem. One of the best articles is Skeel (2001), which focuses on three examples: the California state pension fund’s list of firms with poor governance, the *Business Week* and *Fortune* lists of America’s worst boards of directors, and Robert Monks’s battle to shame the board of Sears into changing the company’s policies. Skeel’s emphasis in all of these is how norms affect the individuals who govern a corporation, and he makes a persuasive case that norms do change their behavior. This is an effective counter to the skepticism of Kahan (2001), who wonders whether the idea of norms has much to add to corporate law unless “norms” is defined to include ideas already used from game theory and other sources. Norms may not be able to explain why corporate law takes the form that it does, or how corporate law should be shaped, but it may be very helpful in explaining how corporations behave within a given framework.

d. Property and intellectual property law. Norms are sometimes considered to be the origin of property, since property can exist in the absence of government, but in our terminology the origin is a convention. Some property rules can be modeled as simple coordination games like that in Table 1 above, for example, the decision of people not to fruitlessly try to use the same land or airwaves at once. In other cases they are in the category of coordination games variously known as Hawk-Dove, Chicken, or the Battle of the Sexes, in which there are two or more equilibrium conventions, but players differ in which convention they prefer. In the Hawk-Dove game, two identical players, Row and Column, contend over a valuable good. Each player chooses whether to be an aggressive
“Hawk” and fight for the resource if necessary, or a timid “Dove” who retreats rather than fights for the good. As Table 3 shows, the best outcome for a player is if he plays Hawk and the other player selects Dove, and the worst is if both players choose to play Hawk, which results in a destructive fight.

Table 3: Hawk-Dove

<table>
<thead>
<tr>
<th></th>
<th>Hawk-Aggressive</th>
<th>Dove-Timid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawk-Aggressive</td>
<td>-1, -1</td>
<td>3,0</td>
</tr>
<tr>
<td>Dove-Timid</td>
<td>0,3</td>
<td>1, 1</td>
</tr>
</tbody>
</table>

Payoffs to: (Row, Column)

The two pure-strategy equilibria are (Hawk, Dove) and (Dove, Hawk). These are asymmetric equilibria, and the question naturally arises of how the players know that the convention is, for example, that Row gets to be the Hawk and Column plays Dove. Without further information, the model cannot answer that question. There does exist a symmetric equilibrium, but it is in mixed strategies. Each player chooses Hawk with probability 2/3 (yielding an expected payoff of (2/3)(-1) + (1/3)(3) = 1/3) and Dove with probability 1/3 (yielding an expected payoff of (2/3)(0) + (1/3)(1) = 1/3). The payoffs in the mixed strategy equilibrium are (1/3, 1/3), below the average payoff of 2/3 that arises if conventions are used.

Maynard-Smith & Parker (1976) noted that both players would prefer a pure-strategy equilibrium if somehow they could take turns playing Hawk and getting the high payoff. One way to take turns is to expand the model and establish a convention that the player who arrives first and claims the valuable good plays Hawk and the second player to arrive plays Dove, “The Bourgeois Strategy.” This is a symmetric equilibrium in the expanded game, and has higher expected payoff than the mixed strategy. Such behavior looks very much like a norm of ownership rights for the first-possessor, but in our terminology it is a convention, being based solely on shared expectations of who will fight and who will flee rather than on any notions of right and wrong independent of material consequences.

Once the convention emerges, it is easy to imagine how it becomes a norm. Among other mechanisms, parents instructing their children to respect the convention – do not take goods first possessed by others and do not let others taking goods you first possessed – would intentionally or unintentionally instill a sense that such unconventional takings were unfair and wrong. In any event, today law and norms of property are largely
co-extensive. Norms against theft and “misuse” support legal property rights, and vice versa. Miller (2003), for example, examines the internalization of the legal norm against parking in “handicapped spaces.”

There are, however, exceptions: norms that constrain the exercise of property rights and norms of “property” that are unsupported by law. First, consider how norms sometimes oppose the exercise of legal property rights. A scholar who sues another scholar for infringing his copyright by photocopying his book chapter may face social penalties. Though orchard owners have the legal right to grow apple trees without keeping bees on their property to pollinate them, and to free-ride off the bees kept on neighboring orchards, Cheung (1973) showed that Washington state orchard owners followed a norm of keeping bees proportionate to their orchard size. Ellickson (1992) famously showed how local Shasta County, California norms governed relations between neighboring ranchers to the exclusion of law. Thus, even when legal rules governing animal trespass damages or the maintenance of boundary fences create certain legal rights, an individual would forgo those rights and follow the norm.

Norms constrain the use of public as well as private property. Ellickson (1996: 1172) thinks of urban problems using the paradigm of “a public space as an open-access territory where users are prone to create negative externalities.” These problems have traditionally been regulated in large part by unwritten rules – either unwritten (or vaguely written) rules enforced entirely according to the discretion of local officials – vagrancy laws, for example – or norms. There is a norm, for example, that a person should not make excessive use of a nonpriced good such as a park bench. Someone who spends the entire day on the park bench with the best view of the White House, whether a vagrant or a journalist, is violating that norm, which establishes a temporal limit to the right to use public property. Ellickson also discusses how norms establish informal zoning for behavior. A typical city dweller drastically changes what he considers bad behavior, worthy of reprimand, depending on where the behavior occurs. He may heap abuse – or at least raise his eyebrows noticeably – on someone inebriated in a residential neighborhood while tolerating much worse inebriation in Skid Row or the Red Light District.

Second, consider how norms sometimes create quasi-property rights, not recognized by law. A scholar who uses someone else’s idea without attribution may avoid violating copyright but may be punished by norms against plagiarism. See Green (2002). We have already mentioned that Epstein (1999) discusses the importance of norms in establishing informal property rights in parking spaces. Sometimes, it is because conventions or norms recognize a quasi-property right that a court will give the right legal recognition outside of statutory law, as when the U.S. Supreme Court recognized a property right in news in International News Service v. Associated Press, 248 U.S. 215
(1918), a case analyzed in Epstein (1992).

Intellectual property presents its own interesting set of issues. Software has proven to be an interesting industry for the study of norms, perhaps because the Internet is new and important enough to have stimulated the creation of new norms. Strahilevitz (2003) focuses on the role of optimistic lies (“charismatic code”) in establishing norms. Gnutella was a network that allowed members to share computer files. It told them, “Almost everyone on Gnutella.Net shares their stuff,” which is false – only one-third of users shared. Also, networks like Gnutella “by no means cede the moral high ground.” despite the dubiousness of their interpretation of the copyright laws. Rather, they try to create norms of cooperation, starting from the general norm of reciprocity. They call people who download files without making their own files available for upload “freeloaders,” even though the record companies might use the same term against Gnutella.

A number of scholars, including McGowan (2001); Benkler (2002) and Lerner & Tirole (2002), have examined the puzzle of why people create open-source software – software that is given away for free, yet is costly to produce. Benkler discusses enterprises such as the Linux operating system, Napster music distribution, and Project Gutenberg – electronic texts that rely on thousands of volunteers to contribute their effort towards a public good. He makes the important point that by dividing the effort into small parts, these enterprises can make do with a small amount of norms; no one person must make a large sacrifice, and each participant can feel satisfaction at having aided the common good.

e. Criminal Law. Criminal law is intimately linked with norms, as one might expect from the fundamental idea of *malum in se* versus *malum prohibitum*. That a crime is *malum in se* – wrong in itself – means that the law prohibits something that also violates a norm, such as theft and unjustified killing. Thus, norms and criminal law may reinforce each other by sanctioning the same conduct. Even if the crime is *malum prohibitum* – wrong only because illegal – the norm of obeying the law may generate some compliance above that predicted by the expected sanction. The level of this effect, however, may depend on the law’s or law-maker’s perceived “legitimacy” (see Thibaut & Walker (1975); Robinson & Darley (1997); Kaplow & Shavell (2000:370)) both because people are more likely to obey a legitimate rule and more likely to cooperate with police in apprehending those who violate legitimate rules. Norms may also help us understand particular crimes. McAdams (1996), for example, claims that norms of privacy help to make the prohibition on blackmail efficient.

Criminal law is the most common outlet for “expressive law”: laws that are meant to express disapproval more than to actually punish it (See Dharmapala & McAdams (2003), Kahan (1998), McAdams (2000), Sunstein (1996)). It might be a crime to commit
adultery or to disrespect one’s parents, but the law’s purpose and effect may be more to express disapproval as to detect and punish these hard-to-prove norm violations. Prosecutorial discretion results in the laws remaining purely expressive; if private law were used, the courts would have to deal with messy civil lawsuits induced by the financial incentives.

The normative economic analysis of criminal law focuses on optimal punishment. Norms are relevant here as well. There is a strong norm of retribution (Kaplow & Shavell 2002: 352-59), which means that there is a taste for punishing wrongdoers and also a taste for the punishment being proportionate to the crime. The presence of this taste has direct and indirect effects on optimal punishment. The direct effect is that punishment does not only deter and incapacitate, but satisfies or dissatisfies tastes for proportionate retribution (which may also in turn affect perceived legitimacy). Thus, where optimal punishment might otherwise be low (as where the risk of detection is high), the norm of retribution may require that it be higher. Conversely, where optimal punishment might otherwise be high (as where the risk of detection is low), the norm of proportionality – that the punishment “fit” the crime – might require that it be lower. See Polinsky & Shavell (2000).

The indirect effect is that the norm of retribution means that some individuals will punish a wrongdoer privately, which affects the optimal level of legal punishment. Optimal deterrence, for example, depends on total sanctions for wrongdoing, not just governmental sanctions. Cooter & Porat (2001) argue that when a tort or contract breach also constitutes a norm violation, it is generally advisable to deduct from tort or civil liability the amount of private sanctions the wrongdoer incurs (and even, in theory, to deduct certain external benefits norms violations create, as where business is diverted to one’s competitors). The same general point may be made about criminal liability. Stigma is an important punishment for wrongdoings, one that combines both public and private punishers. As Rasmusen (1996) explains, the court’s official declaration that someone has committed criminal actions can be important even if there is no material punishment by the state, because the information thereby transmitted makes private actors behave differently towards the criminal. A controversial current application of this idea detailed by Teichman (2004) is in “Megan’s Law” statutes, which publicize the identity and living location of sex offenders. Kahan (1996) uses the differing ability of public sanctions to condemn and stigmatize to explain political support for or opposition to alterative sanctions.

Another effect of non-legal sanctions arises on the issue of optimal sanctions for repeat offenders. In addition to other reasons, Dana (2000) justifies higher legal sanctions for repeat offenders on the ground that the probability of incurring non-legal sanctions declines with the number of violations (because non-legal sanctions such as boycotting
are often exhausted after the first or second violation or because the repetition of offenses signals the fact that the offender is not a member of a community that will subject him to non-legal sanctions for such violations). To maintain the same total sanction, it would then be necessary to raise the legal sanction.

f. Discrimination and Equality Law. Discrimination law is similar to morals law in being closely entangled with norms specific to a time and place. Norms and conventions often govern behavior according to the social groups to which someone belongs, particularly sex, race, ethnicity, or religion. In various times and places, norms or conventions defining sex roles have allocated some jobs exclusively to men and others exclusively to women (Hadfield, 1999); compelled women to take their husband’s surname upon marriage and stay at home to rear children; and differentiated the sexes by dress. Racial, ethnic, and religious group norms often require that members of a group adhere to distinctive codes of dress or food consumption that publicly identify group membership or loyalty. See Kuran (1998). Other norms or conventions compel group members to “discriminate” against non-members, as by prohibitions or limitations on economic transactions or marriages outside the group, the refusal to accord non-group members customary signs of respect, or even the use of violence to suppress non-group members in competition for scarce resources or governmental control. See McAdams (1995) and E. Posner (2001).

A number of scholars have discussed such norms. Akerlof (1980, 1985) claims that “customs” of caste may survive market competition because third parties punish those who violate the custom, though he doesn’t explain why third parties incur such enforcement costs. McAdams (1995) offers to explain third party enforcement by the “payment” of esteem or status. Discrimination arises as groups compete for social status and individual members are rewarded with intra-group status for contributing their group’s societal status by discriminating against others or by punishing non-discriminators. E. Posner (2001) instead describes race discrimination as a convention (in our terms) that emerges from a signalling game. In his model, individuals incur costs to conform to the convention to signal their low discount rates. These authors view discriminatory norms as socially costly (e.g., E. Posner 1996: 1722-23). Even those who are relatively optimistic about group norms have predicted efficiency only where norms primarily affect group members, and have expressed pessimism about the external effects of norms on non-group members– see Ellickson (1991:169) and Cooter (1996:1684-85). Indeed, Kuran (1998) raises the concern of sudden “cascades” in the level of ethnic identification – which he calls “ethnification” – a process that can lead to violence.

In American history, and in other societies today, law has been used to reinforce such norms. In the Jim Crow era of the American South, state and local laws mandated segregation of certain types of public transportation, barred racial minorities from
attending certain schools, and prohibited miscegenation. More recently, laws seek to suppress and undermine discriminatory norms. One step has been to interpret federal constitutional law to invalidate state law requiring discrimination, as in *Brown v. Board of Education’s* (1954) invalidation of formal racial segregation in public schools, and to prohibit other official action based on discriminatory motives. Similarly, McAdams (2000) argues that the First Amendment’s constitutional prohibition on the state “establish[ing]” a religion might be understood as an effort to prevent cascades toward extreme religious conformity. As a second step, federal law now prohibits private discrimination on grounds of sex, race, ethnicity, religion and other such factors, in employment, housing, lending, public accommodations, and other such domains. A third step has been to use law to permit private individuals to favor previously disfavored minorities through “affirmative action,” or to require government agents to do so. A major part of the debate over affirmative action is whether it works ultimately to promote or undermine discriminatory norms.

g. Family Law. Family law is saturated with the influence of norms. Indeed, it is separate from contract law largely because of a longstanding belief that social norms are crucial to how families will be allowed to make use of the courts – that, unlike in commercial contracts, courts ought not to enforce all marital agreements. Instead, the courts should allow social norms to regulate behavior within the family, even behavior that between strangers might be grounds for suit. The motivation was not only to keep courts out of a sphere in which they could not make well-informed decisions, but also to prevent government from aiding agreements in violation of social norms or from intervening in ways that, as Stephen (1873) argued, would weaken marriage norms. He claimed that law could not govern families as well as norms but could have the unintended consequence of damaging norms. As Rasmusen & Stake (1998) notes, the difficulty of customizing legally enforceable marriage agreements has remained, however, even as social norms have weakened and the default definition of marriage has departed radically from the traditional idea of dissolution only for fault.

While there has been attention to economic models of the family in law-and-economics, there has been less attention to norms. One exception is the 1995 article by Elizabeth and Robert Scott, “Parents as Fiduciaries,” which analyzes the legal role of parents as closer to that of fiduciaries such as trustees who act for beneficiaries than of agents who act for principals. A fiduciary incurs legal liability as well as any norm-based penalty for violations of his duty, but norms enter in defining that duty, an example of the “incorporation by reference” that we discussed earlier. Robert Ellickson’s 2005 “Norms of the Household” takes a different approach, focusing not on the family, but on the related situation where more than one person lives in the same residence with the possibility of exit. A “household” is different from a family not only by including mere roommates, but also by excluding traditional marriages (from which exit was difficult)
and single-parent families (because children cannot exit). Ellickson argues that consensus is a desirable method to make decisions in such an organization.

Another example, which shows the possibilities for empirical work in this area, is Margaret Brinig and Steven Nock’s 2003 article, “I Only Want Trust’: Norms, Trust, and Autonomy.” Brinig and Nock examine data on the mental health and other characteristics of divorced couples. They find that marriage break up after a collapse in trust, which might be a failure of either a convention or a norm. Also, divorced men who fail to gain any custody of their children have a significant increase in depression, although remarriage reduces the amount of depression. Brinig and Nock suggest that the depression might arise as a result of punitive social norms triggered by the disgrace of losing custody, but their data does not permit them to test this against the simpler theory that the men miss the company of their children.

**h. Other Public Law.** We have discussed criminal law, discrimination law, and family law, but these are only three of many areas in which the government regulates behavior. Here we briefly discuss the relevance of norms to tax compliance, environmental compliance, driving behavior, and voting.

Norms are important to understanding tax compliance. Posner (2000a) began the discussion of whether strict enforcement of tax laws is a substitute or a complement for norms of tax-paying. Lederman (2003) argues that stricter enforcement of tax laws is actually complementary to norms of legal obedience. Enforcement will increase the number of people who obey the laws for prudential reasons and creating this “critical mass” of taxpayers will create disutility to others if they violate the law. Evidence for her argument is the experiment conducted by the State of Minnesota, which sent a sample of potential taxpayers a letter telling them, truthfully, that most citizens do pay their state income tax. People who received the letter paid more taxes than those who did not. See also Kahan (2002); Murphy (2004). Kirsch (2004) critiques the use of shaming sanctions and “norm management” as an alternative to traditional penalties for tax avoidance, concluding that the problems of such an approach justify only a narrow use of such sanctions.

There is also some literature on how norms matter to compliance with environmental regulations. Vandenbergh (2003) identifies norms that influence corporate environmental compliance. He discusses the empirical evidence for the existence of several relevant norms, including the substantive norms of law compliance, human health protection, environmental protection, and autonomy. He explores the implications of these norms, concluding that future environmental enforcement policies should strive to harness them or at least to avoid undermining them. Carlson (2001) looks at the effort by local governments in the United States to influence individual behavior by inventing new
norms of recycling. She claims that the most important policies are those that reduce the cost of recycling rather than those that try to change people’s preferences, make signalling more effective, or direct esteem towards those who recycle. The question remains, however, why anyone bothers to incur positive costs to recycle. Carlson concludes that signalling and the desire for esteem are important, even though the cost of recycling can easily swamp their effect.

Traffic laws are another fertile area. Sugden (1986; 1998) frequently uses a simple traffic conflict – the “Crossroads Game” – to illustrate his theories of the evolution of conventions and norms. Strahilevitz (2000) provides a case study of traffic compliance that explores the effect of “commodification” on norms, a matter that has worried some theorists. He studies San Diego’s FasTrak carpool lane program. Under FasTrak, drivers could either carpool to be allowed to drive legally in special fast lanes for free or pay a price to drive in the fast lane without carpooling. Establishing a price for the fast lanes for non-carpoolers actually increased the amount of carpooling. In addition, the price was paid by many non-carpoolers who had before been violating the rules by driving alone in the fast lane. Strahilevitz suggests that this is because “[t]he commodification of the road makes other drivers less sympathetic to cheaters. The Express Lanes violator is transformed from a rebel into a scofflaw.” (p.1231).

Finally, voting has posed a challenge for rational choice theories because the expected benefits from influencing an election are so small compared to the costs of voting. Hasen (1996) offers norms as a possible explanation: as with other types of socially beneficial behavior, in some communities individuals receive small social rewards for voting or small sanctions for not voting. In this sense, voting rates may reflect the degree to which a community more generally succeeds in encouraging privately costly but socially beneficial behavior among its members. Because American society has relatively low levels of voting, Hasen explores the possibility of creating a legal duty to vote and supplementing the informal incentives with legal sanctions, as a few European nations do.

i. Constitutional law. A constitution cannot itself be based on law, since law is only established by the constitution, a meta-law. Thus, compliance with the constitution must be based on norms or convention. A variety of scholars, including Hardin (1991) and Buckley & Rasmusen (2000), discuss constitutions – written and unwritten – as particular equilibria of coordination games and consider how norms may help support such constitutions. Writing down the constitution, as in the United States, helps to establish the equilibrium, but the equilibrium does not require writing, as the British Constitution shows. Even where there is a writing, it may be irrelevant to how things actually work (see also Ordeshook (2002)). In addition, there is a normative element to constitutions. Certain government action is thought to be wrong and called “unconstitutional,” a
pejorative term used not just in the United States, but in Britain, where there is no Supreme Court to officially label behavior as unconstitutional. At least one written constitution – the 1793 Constitution of France – enshrines the norm of rebellion: “When the government violates the rights of the people, insurrection is for the people, and for each part of the people, the most sacred of rights and the most indispensable of duties” (“Declaration of the Rights of Man and Citizen,” Article 35 – a provision prudently omitted from the Constitution of 1795).

**j. International Law.** International law is a natural setting in which to expect norms or conventions to be important, because there is no authority above nations to enforce the rules by which they behave. Whether norms actually arise is an open question. Goodman & Jinks (2004) argue that international law can influence states by “acculturation” of state actors. J Goldsmith & Posner (1999) offer a more skeptical view. They argue (at p.1132) that nations generally “do not act in accordance with a norm that they feel obligated to follow; they act because it is in their interest to do so.” Frequently, nations will view it in their interest to comply with their international treaties, which essentially define the parameters of cooperation and defection in an iterated prisoners’ dilemma model. See also Guzman (2002). By contrast, Goldsmith & Posner claim that the non-treaty obligations international lawyers term “customary law” are really no more than the description of what states have found it in their interest to do in the past, which does not even state a convention governing future behavior. Ginsburg & McAdams (2004) also emphasize a state’s self-interest but envision a slightly larger role for international law. They contend that international adjudication generates compliance not because of a norm of complying with legitimate authority but because the adjudication signals disputed facts and clarifies disputed conventions, and in each case it is then often in the parties’ interest to comply.

Under any of the latter three accounts – Goldsmith & Posner (1999), Guzman (2002), and Ginsburg & McAdams (2004) – international law works to the extent it does because it is a convention – under the terminology of this chapter – and not a norm. This is a useful conclusion, if true. It suggests that the existing successes of international law (e.g., rules on diplomatic immunity and on the treatment of neutral shipping during war) have been achieved without internalization of incentives and raises the question of whether international law can achieve further success without first creating a genuine norm.

**6. Conclusion: The State of Research on Norms**

For its first two decades, law and economics largely ignored social norms and conventions. In this period, law and economics scholars implicitly embraced “legal
centralism” – the idea that government provides the only source of order, and law the only set of enforced rules. In the 1990's however, law and economics “discovered” social norms and (in the terminology we use) conventions as sources of what Ellickson (1991) calls “Order Without Law.” Incorporating informal order into the analysis substantially changes both normative analysis and positive predictions of behavior with and without law.

The effects are so varied and pervasive that they are difficult to summarize, but we note a few major points. Where there was once a presumption that a given legal rule influenced behavior, there is now greater appreciation of the need for empirical research to verify a law’s influence, given the possibility that a norm produced the required behavior prior to (and independent of) the law, or that a norm causes people to ignore the law. Where it once appeared that law offered the only solution to the market failure of externalities, we now see that norms often work to punish those who create negative externalities and reward those who create positive ones. At the same time, where theorists once emphasized the need for law to overcome collective action problems caused by individuals maximizing their narrow self-interest, theorists now recognize an important role for law to correct inefficient conventions and norms. Where welfare economics once gave little attention to the fact that the rules it identified as optimal might be perceived as unfair, it is now more accepted that such perceptions are common and their effect on utility – whether or not it fits the taste of the analyst – must be incorporated into social welfare. Where it once seemed that legal compliance was simply a function of deterrence and incapacitation, we can now explain why the norms of legal obedience and the rule of law matter too, and how more specific norms and conventions can either reinforce or undermine legal sanctions. Indeed, the very operation of some core legal institutions – the jury, the police, the bar – may depend significantly upon the norms that regulate them, and we cannot say whether an institution is efficient or inefficient without knowing which norms are interacting with it.

The breadth of the norms literature can also be understood by the variety of issues and legal topics that it has addressed – property, torts, contracts, criminal law, tax, etc. Here is a sign that the literature is maturing. The initial wave of norms scholarship in the early to mid 1990s tended to discuss the general topic of norms and to justify its importance by considering various puzzles or anomalies that could be explained only by norms. Some papers did focus on a narrow legal topic or problem and used norms as one part of the economic toolkit, but it is only in the past few years that this form of scholarship has dominated. We take this change as a sign of progress. Norms are no longer the concern of only “norms scholars” but of a large set of law and economics scholars – indeed, of rational-choice scholars generally – who see norms as one useful concept among many for understanding behavior.
In this regard, the literature retains huge potential, as there are many areas of law in which there has been little or no attention to norms, and since so much of the work will require detailed empirical examination. Indeed, Scott (2000: 1644) turns the tables on the heterodoxy nicely, saying that scholars in “behavioral” and “norms” law and economics alike have fallen into the “Fundamental Attribution Error,” – “the experimentally observed tendency of humans to make the mistake of overestimating the importance of fundamental human traits and underestimating the importance of situation and context.” By this he means that we scholars like broad theories and dislike the hard work of learning the facts of particular situations. This is a valid point, and, indeed, economists fall into this trap when they reject “taste-based” explanations of behavior – which require empirical validation to be useful – in favor of more general explanations based on price changes.

Norms scholarship is much like public choice theory. Both give new insight by asking new questions. Public choice theory asks questions such as, “Are the costs of a law concentrated and the benefits diffused?” and “Is this law difficult to understand for those who would be hurt by it?” Norms scholarship asks questions such as, “Is there a reason why this form of disutility would benefit a society in which it exists?” and “Was there a reason for this norm in the past, even if it is pernicious now?” The two ideas are complementary, as any two big good ideas would be (Geoffrey Miller’s article, “Norms and Rents” is a good example of how they can be combined smoothly). Public choice helps explain why an inefficient norm might exist – it might have been to the advantage of certain concentrated interests to create such a norm. And norms help explain why lobbying groups exist – citizens may feel badly if they fail to aid a lobby that helps them, even though the lobby would probably succeed without any one person’s contribution. Yet both norms and public choice are subdisciplines that claim much, and whose reach sometimes exceeds their grasp.

Norms have explanatory power. They explain why so much behavior seems to be efficient, internalizing externalities even when laws and material self-interest do not constrain behavior. We must beware, however, of simply saying “It’s a norm that’s doing it!” whenever behavior seems puzzling. And we must beware of attributing too much influence to norms even when they do exist. It is clear that people act on their principles, but it is also clear that people will sacrifice one principle for another on occasion, or sacrifice a principle for a taste. Any economist knows full well that if the price of a good rises, the quantity demanded will fall, and a principle is in this respect like any other good. Recall the conclusion of Carlson (2001) that convenience was crucial in determining the amount of recycling, and note the psychology experiments recounted in Jeffrey Rachlinski’s 2000 warning that mundane considerations of instructions and inertia often trump even norms as religious beliefs (in the Darley-Batson “good Samaritan” experiment) or opposition to torture (in the Milgram “electric shock” experiment). We,
like most of those who have thought hard about norms, believe that they are important
and useful in explaining behavior. But it is important also not to forget magnitudes of
incentives, or the need to carefully consider how hard it is to change those magnitudes.
Every one of us has principles – but how many of us are principled?
REFERENCES

Articles and Books


Cameron, Samuel (2001) "The Economic Analysis of Social Customs: The Case of Pre-


**Cases and Statutes**
The 1793 Constitution of France

The 1795 Constitution of France


Titus v. Bradford, 20 A. 517 (Pa. 1890)

Texas & Pacific Railway Company v. Behymer, 189 U.S. 468 (1903)

The T.J. Hooper, 60 F2d 737 (2d Cir. 1932)


Model Penal Code

The Uniform Commercial Code