

FAIRNESS, CHARACTER, AND EFFICIENCY IN FIRMS

ROBERT COOTER[†] & MELVIN A. EISENBERG^{††}

In the middling and inferior stations of life, the road to virtue and that to fortune, to such fortune, at least, as men in such stations can reasonably expect to acquire, are, happily in most cases, very nearly the same. In all the middling and inferior professions, real and solid professional abilities, joined to prudent, just, firm, and temperate conduct, can very seldom fail of success.

—Adam Smith¹

INTRODUCTION

What forces shape the conduct of individuals and firms? Legal scholars traditionally focused on the explicit content of the directions that legal rules give to people. Within the last twenty-five years, however, law-and-economics scholars have shifted the focus from explicit directions to the implicit incentives created by legal rules. Within the last ten years, both lawyers and law-and-economics scholars have turned their attention to the interaction between legal rules and nonlegal norms. This recent discussion mostly concerns the norms of large social groups, such as the general society, merchant communities, ethnic groups, or the corporate community. We call norms of this type *general social norms*. In contrast, this Article concerns norms that are specific to firms: in particular, firm-specific fairness norms.

We advance two theses: First, firm-specific fairness norms typically promote efficiency. Second, firm-specific fairness norms best promote efficiency when supported by reputation effects and when the firm's agents internalize the norms. Internalized norms become part of an agent's character. We will develop the concept of *good agent character*, by which we mean agent character that serves the firm's

[†] Herman Selvin Professor of Law, Boalt Hall, University of California, Berkeley.

^{††} Koret Professor of Law, Boalt Hall, University of California, Berkeley.

We thank Lisa Bernstein and Claire Finkelstein for valuable comments on an earlier draft of this paper.

¹ ADAM SMITH, *THE THEORY OF MORAL SENTIMENTS* I.iii.3.5., at 63 (Liberty Fund 1984) (1790).

profitability by embodying the firm's fairness norms. This neglected topic helps explain how firms work and why they exist.

I. DEFINITIONAL ISSUES

We begin with the issue, what constitutes a norm. A common approach to this issue is to define norms as unofficial standards of conduct. We call this type of definition *standards-based*.

Standards-based definitions are unduly exclusive, because in ordinary language the term "norm" encompasses regularities that are not standards, such as the practice of directors, prior to the 1990s, to exercise only limited diligence.² Violating a standard of behavior typically provokes criticism by others, whereas departing from a mere regularity only prompts surprise.

Standards-based definitions are also unduly inclusive, because they encompass types of standards that are not characterized as norms in ordinary language. Norms are only one kind of standard. Another consists of explicit private rules, such as the rules of organized baseball. Although rules of this sort are regularly enforced and complied with, they are seldom if ever referred to as norms. A third kind of standard consists of cultural values, such as whether social standing is based on merit or lineage, whether there is social approval or disapproval of commerce and merchants, and whether the society's ideology emphasizes individuals or groups. Like explicit rules, cultural values are usually not called norms.

In short, social standards range in a spectrum from explicit private rules, through norms, to cultural values. Two elements characterize movement along this spectrum. The first element is specificity. Usually, explicit private rules are highly specific; norms are less specific; and cultural values are highly general. The second element is the mechanism of creation and change. Explicit private rules are promulgated, amended, and repealed. Norms evolve as practices and commitments shift. Cultural values are transmitted through socialization into a way of life.

Bearing all these considerations in mind, we define norms as *social standards and regularities other than explicit private rules or cultural values*.

Like the general society, every firm has its own rules, norms, and cultural values, which we call *firm-specific*. Our particular interest in

² Melvin A. Eisenberg, *Corporate Law and Social Norms*, 99 COLUM. L. REV. 1253, 1266 (1999).

this Article is in firm-specific fairness norms, that is, norms of a firm that require fair conduct by its agents.

II. FIRM-SPECIFIC FAIRNESS NORMS AND EFFICIENCY

People can often create value by cooperating with one another. The value that they create must be distributed. Disputes over distribution impede cooperation and waste resources. Fairness norms reduce disputes over distribution. Our first thesis is that firm-specific fairness norms promote a firm's efficiency by increasing cooperation among its agents. Next we explain several manifestations of the efficiency of firm-specific fairness norms.

A. *Loyalty*

To begin with, firms can be efficient only if their agents render *loyal* performance, by which we mean that the agents perform their contractual obligations diligently and honestly. Diligent and honest agents create value by cooperating with their firms and co-agents, whereas disloyal agents redistribute wealth to themselves by slacking and deceiving or otherwise betraying their firms and co-agents.

One way to achieve loyal performance is by legal sanctions. Legal sanctions alone, however, have limited effectiveness, because detecting breaches of the duty of loyalty is difficult and legal enforcement is expensive.

Bonding and monitoring provide other methods of achieving loyalty. For example, a corporation may award options to an agent that vest after a fixed number of years of employment. Such options bond the agent to the corporation in two ways. First, because the agent will do well if the corporation does well, the agent has an incentive to do what is best for the corporation. Second, if the corporation detects disloyalty by the agent during the relevant period, it can discharge the agent before the options vest, so that the agent loses the options. Additionally, a firm may employ external and internal auditing systems and various compliance programs to monitor its agents.

Like legal sanctions, bonding and monitoring devices are expensive and have limited effectiveness. A more effective and reliable method for ensuring loyalty is the development of a regime of firm-specific fairness norms. As we explain below, developing such a regime also has costs, but the costs are typically lower than those associated with either legal enforcement or bonding and monitoring.

Legal sanctions, monitoring, and bonding all have an important place in ensuring diligence and honesty by agents, but norms have the central role in achieving that objective.

B. *Supracontractual Performance*

In Part II.A, we explained that firm-specific fairness norms promote efficiency by inducing diligent and honest performance of contractual obligations. However, firms will be most efficient if agents go beyond their contractual obligations. For example, subordinates may put in longer hours and exert more effort than their contract requires, in which case we say that performance is *supracontractual*. Firms have an obvious efficiency interest in inducing supracontractual performance.

A subordinate who performs supracontractually will normally do so because she forms a reasonable, noncontractual expectation that, when the time comes, her superiors will reward her with a raise or a promotion or in some other way. Such expectations are special cases of the well-known phenomenon of implicit contracts. In such cases, however, it is often easy and profitable (at least in the short run) for superiors to exploit the subordinates' trust by not fulfilling the latter's expectations. Accordingly, before agents render supracontractual performance they need a reasonable degree of assurance that their trust will not be exploited opportunistically.

In a world of perfect information and no transaction costs, agents could protect against this kind of opportunism by contract. A perfect contract would encompass the subordinate's future performance, endogenous changes in the firm, exogenous changes that affect the firm, and what constitutes fair treatment under a variety of circumstances. In the real world, contractual protection of this sort usually fails because the needed terms are difficult to specify, observe, and prove.

Firm-specific fairness norms enter to supplement imperfect contracting.³ Two kinds of firm-specific fairness norms are relevant. The first consists of firm-specific versions of general fairness norms, such as "Reciprocate within the firm," and "Don't lie within the firm." These norms are firm-specific rather than general, because at least in principle they are compatible with norms that tolerate acting nonreciprocally and telling lies to people outside the firm. The

³ Benjamin E. Hermalin, *Economics & Corporate Culture*, in THE HANDBOOK OF ORGANIZATIONAL CULTURE AND CLIMATE (S. Cartwright et al., eds. 2001).

second kind of firm-specific fairness norm consists of norms that help define fairness within the firm. For example, compensation depends on seniority in some law firms, while in others it depends on the amount of business that a partner generates. Paying a young partner top compensation could be fair in the second context and unfair in the first.

C. *Mentoring*

Continuity in most firms requires a mentoring system under which superiors train subordinates to assume ever greater responsibility. Mentoring involves significant costs. A good mentor must not only incur time and trouble, but must also share with the protégé valuable skills, techniques, and secrets about the firm and even about himself.

There are various reasons why superiors are willing to incur these costs. Some of these reasons are self-seeking and others are not. In the absence of a regime of fairness norms, superiors are unlikely to mentor subordinates, because the superior will not be able to count on the subordinate to reciprocate appropriately, keep secrets, and so forth.

III. REPUTATION AND GOOD AGENT CHARACTER

We previously explained that firm-specific fairness norms promote efficiency by encouraging loyalty, supracontractual performance, and mentoring. Now we explain why people might respect firm-specific fairness norms and adhere to them. In an effective regime of firm-specific fairness norms, the firm's agents have some assurance that adherence will be substantial and not merely nominal. Two mechanisms especially provide such assurance: reputation effects and good agent character. We explain several ways that reputation and character support fairness norms.

A. *Reputation Effects*

An agent may adhere to firm-specific fairness norms because he believes that acquiring a reputation for adhering will further his own interests. For example, if a superior develops a reputation for truth-telling, reciprocity, and trustworthiness within the firm, he will be more able to induce supracontractual performance from subordinates. Conversely, if a superior develops a reputation for lying, nonreciprocity, and untrustworthiness within the firm, he will

be less able to induce supracontractual performance. Therefore, a superior may tell the truth, reciprocate, and act like a trustworthy person, not authentically, because he has internalized firm-specific fairness norms, but instrumentally, to obtain the reputation that he needs to induce supracontractual performance from subordinates.

As a method of making firm-specific fairness norms effective, reputation is important but imperfect. For example, the existence of an implicit contract and the failure of a superior to honor such a contract are often difficult to demonstrate to third parties. Furthermore, a superior's reputation only imperfectly follows her move to another firm or even her transfer within the firm. A subordinate therefore knows that imperfect information gives superiors significant leeway to break implicit contracts without a loss of reputation. Reputation effects are especially problematic when a relationship is ending, because the two parties no longer have the expectation that they will receive future rewards for conveying present benefits.

A subordinate also knows that if a superior's only motive for acting fairly is reputation, then when the time comes to reciprocate the superior will make an instrumental calculation of costs and benefits. Specifically, the superior will consider whether his gain from failure to reciprocate will exceed his loss of reputation discounted by the prospect of detection. Bruce Chapman explains how these calculations can defeat themselves:

[Suppose that an individual who acts only on an instrumental basis makes a promise, and put aside the issue of legal enforceability.] The question . . . is whether such a promise can be a credible one in the eyes of . . . other individuals. The difficulty is that a promise made for only an instrumental or forward-looking reason, the sort of reason that a rational maximizer of her own preferences would provide, is a promise that should, rationally, be broken when it pays to do so. . . . Having promised, for purely instrumental reasons, to [perform, the individual] will find that these *same instrumental reasons* tell her to break the promise that she has made. Furthermore, all this instrumentally rational behaviour can be predicted perfectly by the other individuals; after all, they too are rational in this way.⁴

Kenneth Arrow makes a related comment about trust: "Trust is . . . extremely efficient; it saves a lot of trouble to have a fair degree of reliance on other people's word. Unfortunately this is not a

⁴ Bruce Chapman, *Trust, Economic Rationality, and the Corporate Fiduciary Obligation*, 43 U. TORONTO L.J. 547, 581 (1993).

commodity which can be bought very easily. If you have to buy it, you already have some doubts about what you've bought."⁵

B. *Good Agent Character*

Another mechanism for making a regime of firm-specific fairness norms effective is good agent character. Before considering this specific concept, we must examine the general concept of good character. Earlier we identified standards-based norms, which supply standards for evaluating behavior. The internalization of standards by actors is closely related to good character. Internalization has two aspects that parallel the preceding distinction between regularities and obligations. First, an actor may adhere to an internalized norm reflexively. The economist Kaushik Basu describes reflexive adherence to a norm:

[Certain norms stop] us from doing certain things or choosing certain options, irrespective of how much utility that thing or option gives us. Thus most individuals would not consider picking another person's wallet in a crowded bus. This they would do not by speculating about the amount the wallet is likely to contain, the chances of getting caught, the severity of the law and so on, but because they consider stealing wallets as something that is *simply not done*.⁶

An actor may also adhere to an internalized norm out of deliberate commitment. In deciding what to do, a sense of commitment to norms receives weight relative to the actor's self-interest. To illustrate, assume that an actor recognizes that adhering to a norm has more costs than benefits to herself. If she adheres to the norm anyway, then her adherence demonstrates her sense of commitment. In general, adherence against self-interest demonstrates commitment to norms.

Self-interest can be defined narrowly to encompass wealth, power, pleasure, prestige, and little else. Under the narrow definition, self-interest often conflicts with morality. Alternatively, self-interest can be defined broadly to encompass moral values, such as truthfulness, integrity, and generosity. Under the broad definition, self-interest seldom conflicts with morality. Thus the distinction between commitment to norms and self-interest is sharp for the narrow

⁵ KENNETH ARROW, THE LIMITS OF ORGANIZATION 23 (1974), *quoted in* Chapman, *supra* note 4, at 581 n.62.

⁶ Kaushik Basu, *Social Norms and the Law*, in 3 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 476, 477 (Peter Newman ed., 1998).

conception and dull for the broad conception.

The premise that people are motivated by normative commitments challenges much economic thinking, which relies on a narrow self that is interested only in wealth, power, pleasure, and prestige. Normative commitment, in contrast, implies a broad conception of the self. The debate about whether understanding law requires a narrow or broad conception of the self seems fertile. In contrast, the debate about whether all behavior is ultimately self-interested seems tautological and relatively sterile. For this reason, we will not debate about whether or not all behavior is ultimately self-interested. Instead, we will develop our argument that understanding law in general, and the law governing firms in particular, requires a broad conception of the self.

Just as violating a social standard provokes criticism from others, so violating an internalized standard provokes self-criticism and guilt. Those who insist that all behavior is self-interested often want to fold guilt into the actor's calculation of the cost of violating a norm. Certainly people want to avoid guilt and other painful feelings. Reducing normative commitment to the avoidance of guilt, however, misconstrues commitment in a way that underestimates its stability.

To see why, consider two ways to avoid guilt. An actor may avoid guilt by changing either her behavior or her feelings. For people who only seek to avoid psychological pain, eliminating guilt by eliminating commitment to the norm is just as good as eliminating guilt by conforming to the norm. For people committed to the norm, however, these two ways of avoiding guilt are unequal. For a committed person, the only acceptable way to avoid guilt is by conforming to the norm. Consequently, a committed person's adherence to a norm is more stable than the adherence of someone who merely avoids painful feelings.

We have explained that internalizing a norm involves a commitment that manifests itself through diminished reflection or choices against narrow self-interest. The internalization of standards relates to *good character*, by which we mean an authentic disposition to adhere to normative standards, either reflexively or on the basis of commitment even when against interest. To have good character, a person need not always adhere to normative standards. It suffices that she has the disposition to do so and typically does so.

Cynics argue that actors conform to normative standards only for reputational reasons, as opposed to reasons of internalization. We reject that argument because it is contradicted not only by experience

but by experiment. In laboratory games conducted by economists and social psychologists, people persistently cooperate more than predicted by reputation effects. For example, people share payoffs with others when doing so is unnecessary because reputation effects are eliminated by anonymity.⁷

Outside the laboratory, the internalization of norms explains some important types of behavior for government and law that baffle theories of narrow self-interest. To illustrate, lawyers in the United States often abandon more lucrative practices to become federal judges, presumably because they enjoy participating in public life and shaping the law to their vision. In other words, judges sacrifice narrow self-interest for the sake of self-expression, including stamping their normative commitments on public life. As another example, the punishment for tax evasion in most countries, discounted by the probability of prosecution and conviction, is small relative to the gain. Whereas economic models of self-interest predict low rates of tax compliance, some countries, like the U.S. and Switzerland, enjoy high rates of tax compliance. Citizens in these countries apparently have internalized a normative commitment to tax compliance.⁸

Economics explains behavior as the confrontation of preferences and opportunities. For economists, preferences are internal values manifested by choices, whereas opportunities are external constraints. Internalizing a normative standard incorporates it into preferences. In contrast, without internalization an obligation remains external and therefore only affects behavior by constraining it. Economists

⁷ See Max H. Bazerman & Margaret A. Neale, *The Role of Fairness Considerations and Relationships in a Judgmental Perspective of Negotiation*, in BARRIERS TO CONFLICT RESOLUTION 86 (Kenneth J. Arrow et al. eds., 1995); Karen Cook & Karen Hegtvædt, *Empirical Evidence of the Sense of Justice*, in THE SENSE OF JUSTICE: BIOLOGICAL FOUNDATIONS OF LAW 187, 197-200 (Roger D. Master & Margaret Gruter eds., 1992); Elizabeth Hoffman et al., *Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology*, 36 ECON. INQUIRY 335, 347 (1998); Elizabeth Hoffman et al., *Preferences, Property Rights and Anonymity in Bargaining Games*, 7 GAMES & ECON. BEHAV. 346, 371-72 (1994); Elizabeth Hoffman & Matthew L. Spitzer, *Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' Concepts of Distributive Justice*, 14 J. LEGAL STUD. 259, 259-60 (1985); Matthew Rabin, *Incorporating Fairness into Game Theory and Economics*, 83 AM. ECON. REV. 1281, 1283 (1993); Paul Graham Loomes Burrows, *The Impact of Fairness on Bargaining Behavior* (1990) (unpublished manuscript, on file with authors). See generally Margaret M. Blair & Lynn A. Stout, *Trust, Trustworthiness, and the Behavioral Foundations of Corporate Law*, 149 U. PA. L. REV. 1735 (2001).

⁸ Eric Posner discusses this phenomenon and provides an alternative rationale in *Law and Social Norms: The Case of Tax Compliance*, 86 VA. L. REV. 1781 (forthcoming 2001).

typically assume that actors regard their reputations as imposing only external constraints, not as expressing internalized values. This view either denies that people have preferences to adhere to normative standards or denies that people's preferences to obey normative standards affect aggregate behavior in equilibrium.

So far we have been discussing *good general character*, by which we mean the disposition of a person to adhere to *society's* normative standards, reflexively or against interest. Good general character, however, is not our main concern. Our main concern is with *good agent character*, by which we mean the disposition of an agent of a firm to adhere to the *firm's* normative standards, reflexively or on the basis of commitment even when against interest.

General and agent character are separable psychologically. To illustrate, the Mafia specializes in illegal businesses that require loyalty among members and ruthlessness towards outsiders. Thus, the Mafia prizes members with good agent character and bad general character. Conversely, the two forms of character converge when an institution requires its members to treat outsiders the same as insiders. Thus, the Catholic Church ideally requires priests to practice the same honesty towards everyone.

Earlier we discussed limitations on reputation as a mechanism to ensure adherence to the firm's norms. Specifically, reputation fails to ensure adherence when an opportunity arises to make a large gain, net of reputational costs, by exploiting another's trust. In these circumstances, a calculation of benefits and costs prompts the actor to sacrifice reputation. Similarly, reputation does not deter unfairness when the relevant community cannot detect that an implicit contract was broken. A person with good character, however, does not behave opportunistically. Instead, such a person passes over such opportunities and reflexively or deliberately chooses against interest. Thus, good agent character generally overcomes the limitations of reputation.

IV. HOW FIRMS ASSURE GOOD AGENT CHARACTER

We conjecture that almost every firm benefits from its agents dealing fairly with itself and one another, although the actual content of fairness norms differs among firms. Some firms also benefit from their agents dealing fairly with outsiders, in which case good agent character goes with good general character. Other firms, however, benefit from their agents dealing unfairly with outsiders, in which case good agent character goes with bad general character.

How can firms ensure good agent character? The most important mechanisms fall into the categories of screening, filtering, education, and socialization. *Screening* occurs when the prospect that a person will have good agent character is taken into account in hiring decisions. *Filtering* occurs when agent character is taken into account in promotion decisions. *Education* occurs when the firm voluntarily undertakes programs, such as compliance programs, that communicate the firm's norms to its agents. *Socialization* occurs when informal interactions communicate the firm's norms among its agents. For example, the firm may encourage agents to think of themselves as firm players, promote social events that increase cohesion among its agents, and so forth.

Screening and filtering require people to perceive the character of others. The obstacle to accurate perception is that persons with bad character may dissemble good character. The ability to dissemble effectively is limited because character is *translucent*; that is, people can see through actions into character, although not perfectly. In statistical terms, translucence means that people correctly infer authentic character from actions with higher frequency than chance would produce. If people could not infer authentic character from actions at a rate higher than chance, then rational people would never attempt to make such inferences. If no one attempted such inferences, then no one could deceive anyone about character. The very possibility of deceit implies both that many people assimilate normative standards into their character and that people infer authentic character from objective evidence with less error than chance would produce.

In general, nature creates a parasite for every host. The very existence of a parasite presupposes the host's existence. Dissemblers are parasitical not only because they redistribute rather than create, but also because dissembling by some implies internalization and authenticity by many.

Like reputation, good agent character only imperfectly ensures adherence to firm-specific fairness norms. But just as imperfection does not preclude an important role for reputation effects in firms, so the possibility of deception does not preclude an important role for character in firms.

In the next Part we discuss processes through which a person can acquire good character. Firms seldom provide the means by which people acquire good character; rather, individuals with good character typically bring it with them to firms. Therefore, screening

and filtering dominate education and socialization in the employment practices of most firms. If firms reward good agent character, however, then persons without good agent character have an incentive to change. The next Part discusses how people change their character.

V. THE ORIGINS OF VALUES

People often change their characters to improve their opportunities. For example, bankers may cultivate trustworthiness and computer scientists may cultivate creativity. Similarly, agents of firms can often increase their opportunities by acquiring good agent character. Superiors with good agent character have an advantage over others in inducing supracontractual performance. Superiors who induce supracontractual performance from their subordinates increase their probability of success leading to promotion. Similarly, subordinates with good agent character have an advantage over others in gaining trust from superiors. Trusted subordinates have a higher probability of distinguishing themselves and securing promotion.

Superiors also have an incentive to develop more insight into the character of others, since more insight enables them to screen, filter, educate, and socialize more effectively. Correspondingly, insightful subordinates are more likely to attach themselves to superiors with good firm character, who will repay their trust. Good agent character in a subordinate also increases the probability of being mentored by a superior. Insightful superiors will tend to adopt protégés who will repay their trust. In short, while agents with bad character are parasitic, agents with good character and agents with good insight into character are symbiotic.

Agents with authentic good character and agents who are good readers of character enjoy a competitive advantage in finding partners for cooperative ventures. Evolution has, consequently, fitted us to signal and detect character. Thus, evolutionary forces operating among and within firms should select for both good agent character and good insight into character.

Insofar as people with good agent character enjoy better careers, agents with bad agent character have incentives for improvement. Improving character involves internalizing normative standards and thereby adopting them as preferences. While the fact of internalization is easily demonstrated, explaining how internalization comes about is much more difficult. Changing one's preferences requires technique. The plethora of self-help books and therapists

testify to the fact that one cannot automatically become who one wants to be. Psychologists have extensively researched the internalization of norms. Notably, Piaget and Kohlberg have studied stages in the development of moral reasoning among children.⁹ According to their theories, children perfect the ability to internalize moral norms as they acquire a capacity for general reasoning. This research makes the process sound rational and cool.

In contrast, "depth psychology" often traces the internalization of morality to irrational processes that are inchoate and hot. According to these theories, internalization of morality ingrains new impulses in a child through emotional experiences. An example is Freud's theory that morality is the "ghost in the nursery," meaning the repressed memory of parental punishments.¹⁰ Repression transmutes fear into guilt. The Freudian idea that internalization attaches a "guilt penalty" to violating a norm fits well with economic models in which behavior flows from a calculus of psychological benefits and costs.¹¹ More recently, theorists have argued that people experience "dissonance" when facts conflict with commitments. To relieve dissonance, people change their beliefs about the facts or their commitment to the values.¹²

⁹ Piaget presented his ideas about stages in mental development in a series of books written in French beginning in early 1930s, including JEAN PIAGET, *THE MORAL JUDGMENT OF THE CHILD* 13-108 (Marjorie Gabain trans., 1965) (1932). Kohlberg also developed his ideas in a series of books and articles over many years; see Lawrence Kohlberg, *The Philosophy of Moral Development: Moral Stages and the Idea of Justice*, in 1 *ESSAYS ON MORAL DEVELOPMENT* 1, 409-12 (1981), in which the appendix outlines his account of the six stages of moral development. Flaws in Kohlberg's approach have generated much criticism from feminists, notably CAROL GILLIGAN, *IN A DIFFERENT VOICE: PSYCHOLOGICAL THEORY AND WOMEN'S DEVELOPMENT* 18-23 (1982). See also Nona P. Lyons, *Two Perspectives: On Self, Relationships, and Morality*, in *MAPPING THE MORAL DOMAIN* 21 (Carol Gilligan et al. eds., 1988).

¹⁰ In Freud's account, morality is the repressed memory of punishment and threats from a child's father. In technical terms, the superego emerges when a child represses her Oedipal fears and identifies with her father. See SIGMUND FREUD, *THE EGO AND THE ID* 18-29 (James Strachey ed. & Joan Riviere trans., W.W. Norton & Co. 1962) (1960). A clear explanation is in RICHARD WOLLHEIM, *FREUD* 177-218 (1971).

¹¹ On the use of a guilt penalty to change the payoff matrix in a game, see MARK CASSON, *THE ECONOMICS OF BUSINESS CULTURE: GAME THEORY, TRANSACTION COSTS, AND ECONOMIC PERFORMANCE* 29-52 (1991). Antiutilitarian philosophers typically reject the theory that conforming to a principle of morality involves weighing alternative reasons and balancing them. For example, see the account of exclusionary reasons in JOSEPH RAZ, *THE MORALITY OF FREEDOM* 267-87 (1986).

¹² G.A. AKERLOF, *THE ECONOMIC CONSEQUENCES OF COGNITIVE DISSONANCE* (1986); Matthew Rabin, *Cognitive Dissonance and Social Change*, 23 J. ECON. BEHAV. & ORG. 177, 178 (1994).

We have mentioned some psychological theories of how preferences change. Perhaps these theories are primitive or perhaps they hold the promise of yielding effective techniques to change oneself. In either case, it is sufficient for most purposes of this paper to assume that people can change their preferences at some cost without explaining how they do it. We turn now to the logic of choosing one's preferences.

Assuming that preferences influence opportunities, how would a rational person choose her own preferences? One possibility is to invoke meta-preferences or higher-order preferences.¹³ To illustrate, a person's choices might reveal a preference for milk over Coke. Behind this choice might lie a preference for health over infirmity. In this example, the first-order preference is for milk and the second-order preference is for health. Behind our particular preferences lie more general, abstract preferences. At the highest level, perhaps some people order all of their specific choices with respect to a supreme value. Traditional candidates for a supreme value among philosophers include pleasure and happiness, whereas economists typically favor wealth, political theorists sometimes favor power, and theologians sometimes favor piety.

The existence of different orders of preferences provides a potential explanation for how a rational person would choose her own preferences. If a person's lower-order preferences affect her opportunities, then she should choose her lower-order preferences so that the resulting opportunities maximize the satisfaction of her higher-order preferences. To illustrate, a eudaemonist would choose honesty over dishonesty provided that honesty increased her happiness. Similarly, an ambitious politician might choose dishonesty over honesty provided that dishonesty increased her power. If the quote from Adam Smith that begins this Article is right, a person in the "middling and inferior stations of life" who wants wealth will cultivate virtue, specifically "prudent, just, firm, and temperate conduct."

In general, higher-order preferences can provide a guide for changing lower-order preferences. Moral philosophers devote much time and energy to understanding rationality in higher-order preferences, whereas ordinary people give relatively little thought to this problem. Many people remain uncertain about their higher-

¹³ Kelvin J. Lancaster, *A New Approach to Consumer Theory*, 74 J. POL. ECON. 132 (1966).

order preferences. To illustrate, many people are unsure about the extent to which they prefer wealth over pleasure, happiness over fame, or virtue over status.

Uncertainty makes higher-order preferences an imperfect guide to choice. Fortunately, people often can make rational choices about their lower-order preferences with little guidance from higher-order preferences. This is possible when the effect of lower-order preferences on opportunities is strong enough. To illustrate, if learning diligence in school promotes both happiness and wealth, then a student who remains unsure about the relative importance of happiness and wealth still has sufficient reason to learn diligence. Or, if Adam Smith is right, a middle class person who remains unsure about the relative importance of virtue and fortune still has sufficient reason to learn prudence and justice.

An idea borrowed from welfare economics may prove powerful in generalizing this insight. Assume that a person with preferences U_o enjoys opportunities F_o . Let X_o denote the point in F_o that maximizes U_o . Also assume that changing her performance in a way that improves her character will improve her opportunities. Specifically, a person with preferences U_o who changes to U_i causes opportunities to change from F_o to F_i . Let X_i denote the point in F_i that maximizes U_i .

Should the person make the change? If the increase in opportunities is large enough, then X_i will be preferable to X_o by either set of preferences, U_o or U_i . In notation, $U_i(X_i) > U_i(X_o)$ and $U_o(X_i) > U_o(X_o)$. After making the change in preferences, the opportunities enjoyed by the person are better than before relative to both her new preferences and her old preferences. This fact provides a reason to make the change in preferences.

A change that is better with respect to the preferences of everyone affected by it is called a Pareto improvement. By analogy, a change in preferences that leaves the person better off with respect to both her original preferences and her final preferences can be called a *Pareto self-improvement*.¹¹

¹¹ This concept is analyzed graphically in Robert Cooter, *Models of Morality in Law and Economics: Self-Control and Self-Improvement for the "Bad Man" of Holmes*, 78 B.U. L. REV. 903, 923 fig.4 (1998). In proceeding from the simple to the complex, microeconomics textbooks first explain the logic of preferences in a timeless world and then introduce the complications of time later. Similarly, we discuss the logical problems of changing preferences in a timeless world. In this Article we have not considered the problem of time-consistency, which requires reconciling rational choice with shifting moods and temporary emotions.

We have explained that the impetus to improve character often comes from recognizing that better character will lead to more opportunities. By offering more opportunities to people with better agent character, firms motivate people to acquire better agent character. Thus firms do not change people so much as firms give people incentives to change themselves.

CONCLUSION

We have explained that agency problems beset firms and prompt opportunistic behavior by employees. Opportunistic behavior redistributes value, whereas cooperative behavior creates value and thus contributes to profits. Agents who have internalized firm-specific fairness norms are less inclined to opportunism and more able to cooperate. Agents who internalize firm-specific fairness norms acquire good agent character and act authentically from internalized values, whereas agents with bad character act instrumentally and opportunistically.

Good agent character conveys an advantage to superiors and subordinates in forming cooperative relations with other people who can read character. Evolution has, consequently, fitted us to signal and detect character. For every host, however, evolution creates a parasite. Effective dissemblers who transmit false signals also enjoy the strategic advantage of deceiving others. Dissemblers, however, are parasitic not only because they redistribute rather than create, but also because they cannot successfully deceive anyone unless other people behave authentically.

We conjecture that almost every firm benefits from effective fairness norms within the firm, although the actual content of these norms differs among firms. Thus, evolutionary forces operating both among and within firms should reward, and therefore select for, good firm character and good insight into character. Firms that survive, however, may deal with outsiders in a variety of ways, some relying on honesty and others on deception. In competitive equilibrium, all firms that survive are equally profitable. The surviving mixture of honesty and deception depends on the nature of the market. Some markets, like insurance, presumably equilibrate at a mixture favoring honest sellers, whereas other markets, like used cars, may equilibrate at a mixture favoring dishonest sellers. Repeat transactions tend to favor honesty, whereas large-valued one-shot transactions may tend to favor dishonesty.

A similar argument applies to individuals. Individuals gain an

advantage from insight into the character of others. Thus evolutionary forces should reward and select for good insight into character. Insofar as people have insight into character, people with good agent character have the advantage of better opportunities to get others to cooperate. Dissemblers, however, have the advantage of eliciting trust that gives them the power to exploit others. Business competition can exert strong pressures on character. Competition tends to equilibrate rates of return. In evolutionary equilibrium, authenticity and dissembling are equally profitable on average.¹⁵ The extent of authenticity and dissembling depends on the characteristics of the market. Authentic signaling is easier to distinguish from dissembling in repeat transactions than in one-shot transactions. In repeat transactions, a person of good character will receive a wage premium from employers who read character well. Consequently, markets with repeat transactions tend to equilibrate when authentic actors greatly outnumber dissemblers.

Distinguishing firms from markets preoccupies the theory of the firm. Our reflections on good agent character suggest a difference between firms and markets. Incomplete contracts aggravate agency problems whose control depends especially on reputation and character. People easily develop loyalty, including obligations of fairness, towards organizations and co-agents. The firm overcomes agency problems partly by developing loyalty among its agents. In contrast, people seldom develop loyalty to markets. To illustrate, an employee of General Motors is far more likely to feel loyalty towards her company than towards the automobile industry. Firms exist partly because good agent character solves the problem of cooperation within firms better than good general character solves the problem of cooperation in markets.

¹⁵ We assume that learning authenticity or dissembling requires a similar investment of resources by the individual.

* * * * *