Inside the Coasean Firm: Competence as a Random Variable

Richard A. Epstein
dangelolawlib+richardepstein@gmail.com

Follow this and additional works at: https://chicagounbound.uchicago.edu/law_and_economics

Part of the Law Commons

Recommended Citation

This Working Paper is brought to you for free and open access by the Coase-Sandor Institute for Law and Economics at Chicago Unbound. It has been accepted for inclusion in Coase-Sandor Working Paper Series in Law and Economics by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
Inside the Coasean Firm:
Competence as a Random Variable

Richard A. Epstein

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

March 2010

This paper can be downloaded without charge at:
and at the Social Science Research Network Electronic Paper Collection.
INSIDE THE COASEAN FIRM:
COMPETENCE AS A RANDOM VARIABLE

Richard A. Epstein

ABSTRACT

The work of Ronald Coase is notable for two primary reasons. First, it introduced the notion of transaction costs to explain the formation and maintenance of firms, Second, it advanced our understanding of the critical topic of social costs. Yet, while transaction costs are key to understanding why firms are organized, they do not offer a complete explanation of how they are organized. A richer account of the problem properly stresses that differences in individual levels of competence, as well as individual variations in temperament and taste, explain why, for example, some firms are organized as partnerships and others as straight employment arrangements, with many permutations in between. Understanding differential levels of competence also helps to explain issues in areas from employment discrimination law to capital markets and tort liability.

INTRODUCTION

In his December 9, 1991 Nobel Memorial Lecture, “The Institutional Structure of Production,” 1 Ronald H. Coase outlined the formative ideas that led him to write his 1937 article, “The Nature of the Firm,” 2 and a generation later, his 1960 article, “The Problem of Social Cost.” 3 In the first of these articles, Coase asks the disarmingly simple question of why, if the price system is so efficient at organizing the relationship between parties, we have any need for firms at all. His answer was, and emphatically is, that it is false to assume that transaction costs are zero in competitive markets. Once the positive costs of exchange are factored into the equation, the difficulties of organizing voluntary exchanges of goods and services in a spot market—including the difficulties in setting price—become evident.

* James Parker Hall Distinguished Service Professor of Law, The University of Chicago; the Peter and Kirsten Bedford Senior Fellow, The Hoover Institution; visiting professor, New York University Law School. This article is prepared for the Conference, Markets, Firms and Property Rights: A Celebration of the Research of Ronald Coase, The University of Chicago Law School, December 4 and 5 (2009). Preliminary Draft. My thanks to the conference participants for their valuable comments and to Sharon Yecies, University of Chicago Law School, Class of 2011 for her usual excellent research assistance.

1 (Coase, 1992)
2 (Coase, 1937)
3 (Coase, 1960)
Recognizing the positive costs of exchange, individuals ask whether the coordination of efforts through the mechanism of a firm is a more efficient way to organize means of production than spot transactions. Where it is, firms will survive in a competitive market. Where it is not, firms will lose out, and other forms of organization will dominate. Any complete analysis of the emergence, persistence and destruction of firms, therefore, requires a close comparison of the costs of the two (or more) possible modes of organization. That analysis, which starts with the archetypical firm, may easily yield intermediate cases, including long-term joint ventures or various types of relational contracts, which exhibit some but not all the characteristics of a fully integrated firm.

In “The Problem of Social Cost,” Coase asked the question of whether Pigovian taxes would be necessary to control pollution externalities in a world of zero transaction costs. His surprising answer—which had to overcome much early (if short-lived) resistance—is that such externalities (including negative externalities like pollution) would never survive in a zero transaction cost world. In all cases, actors would seamlessly correct for the supposed externality by entering into a series of voluntary exchanges, irrespective of the original assignment of the right to pollute or be free of pollution. The actual disposition of rights and the number of affected parties would not matter because an infinite number of transactions could correct all initial imbalances in an infinitesimal length of time.

All the action, therefore, arises in a world with positive transaction costs, where the key challenge of institutional design is to design a system of initial property rights with two desirable characteristics. The first is a clear delineation of rights that permits the efficient use of various resources. The second is a distribution that allows for the transference of property rights so as to facilitate cooperation and exchange. Common law rules on property rights—which give the owner of each asset the exclusive right to possess, use and dispose of a given thing—go a long way toward achieving this goal.

To see why this constellation of common-law entitlements is necessary, one need only ask what happens to a resource when one or more “sticks” is eliminated from the bundle of common law property rights. Possession without the rights of use and disposition is a burden, not a benefit. The possessor is liable for harms

---

4 Pigovian taxes are taxes that force producers to internalize the cost of negative externalities they produce. Coase’s example is “a tax on the factory owner varying with the amount of smoke produced and equivalent in money terms to the damage it would cause.” (Coase, 1960, 1)

5 “In the course of two hours of argument the vote went from twenty against and one for Coase to twenty-one for Coase.” (Stigler, 1988, 76)

6 Property denotes “the group of rights inhering in the citizen’s relation to the physical thing, as the right to possess, use and dispose of it.” United States v. General Motors, 323 U.S. 3777 (1945). Note the reference to citizen is no part of the definition of property.
caused by natural or artificial conditions on the land, but receives no gain from its use. Abandonment of the property would be an owner’s preferred course of action. Possession with the right of use, but not disposition, improves the situation by allowing for use, but precludes gains from trade and cooperation. The right of possession and disposition, but not use, makes it possible to sell property to another person for whom it is just as useless as the previous owner. Only the three incidents of ownership, taken together, yield the correct result because they reduce the transaction costs needed to leave or get the resource into the hands of its most efficient owner. Those rules thus dovetail neatly into the Coasean world view, which has both normative and positive implications: if a set of legal rules can minimize transaction costs, it can also maximize social welfare. (However, Coase has long been reluctant to push his fundamental insights in such a general form.)

It is instructive to note some of the insights that are omitted in this brief exposition the Coasean world view. For example, Coase understands that self-interest is often a guiding force in commercial exchanges, but he does not insist that self-interest necessarily eliminates all other forms of human motivation. He tends to regard strategic behavior among honorable traders as less of a peril than other commentators, which is evident in his long debate with Benjamin Klein over General Motors acquisition of the Fisher Body Company in the mid-1920s. Nor does Coase insist that all individuals are unerringly able to maximize their individual welfare by choosing the means appropriate for the achievement of their ends. These elements are surely a part of the Coasean world view, but they do not dominate it. He would never say that individuals are not capable of generous impulses or charitable actions. On that regard, Coase’s views of human nature owe as much to the Adam Smith of A Theory of Moral Sentiments as they do to the Adam Smith of The Wealth of Nations.

When Coase gave, of all things, the Coase Lecture at the University of Chicago in 2003, he recapitulated many of the points contained in his Nobel Prize Lecture a dozen years before. At its closing, he was challenged by a student who wanted to know whether his view of human nature aligned him with the new devotees of behavioral economics, with its strong emphasis on the cognitive barriers to rational choice behavior, based on such notions as the availability heuristic or the tendency toward anchoring (which cut, of course, in opposite directions). Coase’s answer was that he had no such specific and unalterable human deficits in mind. His view, rather, was that most people are highly imperfect beings who try to do the right thing most of the time, but nonetheless often fail. Given that condition, the purpose

---

7 (See Coase, 2000; Coase, 2006. For the alternative account, see Klein, et. al. 1978; Klein, 1991.)
8 (Coase, 2003)
of a sound legal framework, especially one with clearly defined property rights, is to make it easier for humans to navigate through the shoals, so that over time small individual gains would coalesce into more substantial social gains. On this matter, he is close to Friedrich Hayek, with whom he overlapped at the London School of Economics. By the same token, the perfectibility of Milton Friedman is not part of his lexicon. He disdains any grand view that celebrates the inability of individuals to make rational choices just as he disdains any view that insists that individuals always make such choices, come hell or high water.

In light of this background, one way to appreciate the Coasean approach is to note that he does not conceive of the firm as a black box that operates mysteriously in a marketplace. Rather, it is a complex set of interpersonal arrangements that allows individuals to coordinate their activities in what Oliver Williamson—one of Coase’s most prominent disciples—now calls a hierarchical fashion. One clear implication is that it is dangerous to model competition as a state of affairs in which, in equilibrium, all firms use the same production function to produce goods at the same price. To be sure, the conventional analysis takes into account that different firms could have access to identical assets but only at different prices, so that we could predict which firm is likely to close down first if demand for good shrunk. But the larger truth is that in many markets, firms that produce the same finished goods use radically different methods to get to that point. Some start high tech, while others start low tech and use high tech only when they discover the gaps in their business. Countless other variations matter as well. It is only when one realizes the weight of transaction costs is it possible to see why key issues of management all begin where the black box model of the firm leaves off. Managers have as their job to make the key interstitial decisions precisely because firms arise when it is not possible to establish perfect contingent-state contracts. There is little question that Coase’s one simple insight on the role of transaction costs in firm design offers lasting insights into the operation of market institutions.

**THE NATURAL VARIATION AMONG HUMAN BEINGS**

There is, however, at least one critical assumption of economic theory that Coase did not address that is important for understanding how firms and individuals operate within a marketplace. I refer here to the differences among individuals on matters of competence and taste—which for these purposes I lump together under the heading of "competence"—that they bring to any market setting. Competence continues to influence behavior as individual participate in various types of business transactions either as lone actors or firm members. This is

---

9 (Williamson, 1973)
relevant here simply because much of the diversity that Coase and others observe in the operation of firms is made more pronounced because it is attributable to the differences in the abilities of individuals who participate in market setting. Coase ignores these issues in “The Nature of the Firm” when he simply posits that “an entrepreneur” will organize the factors of production when the transaction costs of running a price system become too great. At no point, however, does he attempt to answer the question of what constellation of individual abilities is conducive to the entrepreneurial role and which not. From a regulatory process, there is no harm in this position, for self-selection will tend to match individuals with their roles. As a practical matter, however, that process could be hastened with greater self-knowledge of the fit between individual talents and particular roles.

More importantly for this paper, the initial distribution of talents on all sorts of dimensions offers an explanation as to why the level of specialization within firms and industries is so great. Because parties are, biologically and socially speaking, not identical in a wide range of relevant characteristics, the gains to trade from specialization are likely to prove larger and more enduring than they would be if the only explanation for the differentiation in occupational roles was based on the relative transaction costs of spot transactions and long-term relational contracts.

As important as these elements are, Coase appears to have had little professional interest: the role of human biology in shaping the psychology of individuals who eventually engage not only in market transactions, but in charitable work, marriage, hobbies, clubs, and civic work. The usual unstated postulate in economics is that competence is fixed and constant across individuals. This point of view, moreover, holds even for those who do not think that all individuals are always rational actors. Thus the extensive and instructive literature on “bounded rationality” “refers to the rate and storage limits on the capacity of individuals to receive, store, retrieve, and process information without error.”¹⁰ Neither this particular quotation nor the surrounding passages give much clue as to whether these limitations are constant across individuals, or what implications flow once we recognize the obvious point that these abilities can vary by orders of magnitude. Behavioral economics also speaks with the same overconfident level of generality in listing the various intellectual infirmities that impede individuals from making rational choices. But once again the key point is that any recognition of these multiple cognitive problems does not carry with it any inference that all individuals suffer from the same impairments, or, if they do, to the same degree. Some persons could be naturally more astute, perhaps because of their greater facility with mathematics, or an ample dose of horse sense. In other cases competence levels could vary with age and experience.

¹⁰ (Williamson, 1973, 317)
Here is one example. In an instructive study, Sumit Agarwal, John Driscoll, Xavier Gabaix and David Laibson\textsuperscript{11} investigate the relationship between age and competence in financial markets. They found variations in borrowing behavior that conform nicely to common sense perceptions of the relevance of age to performance. Although it may not be said allowed in polite company today, people appeal to the common sense notion that on most tasks people first get stronger with age and then gradually decline. That is why everyone speaks about athletes (or scholars) in or past their prime. These age dependent notions, however, fit only with difficulty into strong neoclassical formulations of uniform rationality. Their head note says it all: “The sophistication of financial decisions varies with age: middle-aged individuals borrow at lower interest-rates and pay fewer fees compared to both younger and older adults.” Fifty-three turns out to be the age of peak efficiency. If all this is true, one can see why financial responsibility varies with age. Further work could of course seek to accept the notion of variability but find different patterns for different subgroups. Thus financial traders with huge demands on their time are usually on the younger side, given the level of stress and the need for facility with higher level mathematical models. What matters for the general analysis is not the particular relation between age and competence, but the explicit awareness on the part of firm managers and employees alike that this relationship matters, and need not be constant across different occupational classifications. That knowledge, once acquired, will drive the internal organization of the firm.

The same observation can be made about the self-interest component of rational choice theory. It may well be that all individuals do not act solely in accordance with only selfish motivations. But there is no reason to believe that all individuals are equally generous or equally selfish. The assumption that these traits vary, perhaps in a normal distribution, is again a key element in any accurate overall description of human beings.

Variation on these three dimensions—calculation, bias, and generosity—is important because it tells us much about how individuals sort themselves out in both the firm and the marketplace. The different roles that people choose are not solely left to socialization and opportunity (such as that afforded by a good education). It is foolish to deny the force of any of these social elements, but it is equally foolish to assume that these social factors are the sole source of human differences in a wide range of social settings. Rather, it is appropriate to take a cue from evolution, where natural variation within a population is regarded as a norm to which there are, in practice, practically no exceptions.

Accordingly, all large populations have some distinctive distribution (which may or may not be captured in a normal distribution) over a full range of traits.

\textsuperscript{11} (Agarwal, et al., 2007, hereinafter \textit{Age of Reason})
These distributions range from height and weight (which themselves will vary in predictable fashion with age), to matters of intellect and temperament (which may be subject to similar age variations). These latter elements may be harder to measure and classify, but they are always relevant and are frequently evaluated by standardized psychological tests (such as Myers-Briggs12) which measure the degree to which different persons are extroverts/introverts and the like with an eye to matching personality types with the occupations at which a person is likely to succeed. These tests of course assume that individuals who fall into one of the 16 overlapping personality types will not have the same career path as those which fall into any one of the 15 others. And the same can be said with the overall level of competence in either general intelligence, or particular aptitude for certain kinds of tasks, as when some people are better verbally while others are better quantitatively. The analysis of fits for particular jobs will also include other variables, including the ability to tackle various means-ends relationships and the individual comfort level with pervasive risk and uncertainty.

It is critical, therefore, to take these various traits and to show briefly how they offer an additional tool to explain not only the division of labor in the marketplace, but also indicate which individuals will take what kind of roles and why. The point of this little mental exercise is not to falsify any fundamental economic proposition that celebrates the gains from trade. Nor is it to offer a frontal challenge to either of Coase’s great papers. Rather, it is to identify the additional gains from trade that are achievable pretty much across the board precisely because individuals start with inborn differences that could easily be magnified as they move through life. Quite simply, just as people with different skills stand to gain more from trade than those whose skill sets are the same, so too individuals with different temperaments and abilities are likely to gain more through trade than persons who have no such differences to exploit.

**Specialization within Firms**

In order to see how this analysis works, start with the simplest question about the organization of a two-person firm. The Coasean insight seeks to explain why some firms persist, but it offers relatively little insight into the details of the internal structure of different firms. This topic concerns any lawyer whose job is to help individuals form a common venture. Coase himself was largely indifferent to this question, when he wrote that “those, for instance, Professor Knight, who make mode of payment the distinguishing mark of the firm—fixed income being

---

12 For those keen on taking the test, see http://www.humanmetrics.com/cgi-win/JTypes2.asp.
guaranteed to some of those engaged in product by a person who takes the residual and fluctuating income—would appear to be introduce a point which is irrelevant to the problem we are considering.”13 But mode of payment is relevant to the vital management question, precisely because the division of the spoils is one of the key way in which firms seek to coordinate the behavior of their members and employees.

To see why, stick for the moment with the simplest business choice that two individuals will have to make: whether to form a partnership or an employment relationship. In both, the individuals will combine their labor in a long-term relationship that is decidedly different from the spot contracts that take place in sales markets. The point is disputed by Alchian and Demsetz who reject Coase’s view that employers exert some special power or authority over their employees above and beyond found in “the conventional market.”14 To them employees and purchasers are cut from the same cloth. Thus the employer “can fire or sue, just as I can fire my grocer by stopping purchases from him or sue him for delivering faulty goods…To speak of managing, directing, or assigning workers to various tasks is a deceptive way of noting that the employer continually is involved in renegotiation of contracts on terms that must be acceptable to both parties.”15

The hyperrealist assessment is correct insofar as it notes that the threat to quit or fire in an employment relationship means that parties do make constant readjustments within the framework of an overall contract.16 But the point is still not right because on the Alchian/Demsetz view there would be no reason to worry about the classification of a contract as one of sale versus employment, and even less of a reason to worry about the relationship being classified as one between two partners instead of one between an employer and employee. Yet is painfully clear from both a social and legal point of view that it is not accurate to describe switching your goods and services as “firing” the grocer, let alone for insubordination. The decisions on what items to buy and what not are discrete. The smart grocer that has established a price scale is quite happy to have a customer who switches from one good to another, so song long as the margin of profit is constant across goods. Indeed avoiding transaction costs is imperative, which is why no large business ever lets its customers negotiate prices with clerks on a check out line. Take-it-or-leave-it offers are the only way to process the large volume of small transactions that characterize modern retail establishments. Just think of the impatience of the next person in line.

13 (Coase, 1937, )
14 (See Alchian and Demsetz, 1972, 777)
15 (Alchian and Demsetz, 1972, 777)
16 (See Epstein, 1984)
In contrast, firing an employee results in the disruption of a long-term relationship in which the employer gives orders which the employee, as part of the antecedent, obeys. The difference between quitting and being fired for insubordination is one that matters even if it turns out that no lawsuit is filed in either case. And for our purposes, the Alchian and Demsetz realist account cannot explain why, from the stock of standard legal relationships, two people sometimes choose the partner relationship and sometimes choose an employer/employee relationship. The distribution of control over the venture turns critically on that choice. Partnerships may well be dissolved at will, but the two parties do not simply go their separate ways, as in the case of the customer and the grocer. They have assets that are jointly owned and managed, and the question of how to either sell the assets for cash or to divide them between the parties requires a good deal of thought, which is heavily influenced by the transaction costs involved in the alternative forms of separation.

Thus in some instances, the assets are sold and the cash is divided. In other cases, one party can buy out the other for a price that can be fixed by arbitration, based on book value or agreed to after extensive negotiations. Introduce three parties and it is critical whether the buyout is done by the firm (at which point the two partners remain equal) or done by the partners separately in different proportions, at which point the buyout brings about a shift of firm control.

Answering these multiple questions about firm structure, as the business moves from formation to dissolution, is far more difficult than appears precisely because the differences in partnerships often must into account the relative competence of the two parties. From Roman times on forward, the background norm with respect to a two‐party partnership was that the partners share and share alike on both benefits and losses, unless there is some different allocation of either capital or control rights within the firm, which in turn would govern. But if we stick to the simplest pro rata partnership relative to the simplest employment relationships, the rules are different. The partners decide their respective shares by agreement. The employer gives orders to the employee, which are normally accepted and executed. In the partnership, the gains are divided equally between the parties. In the simplest employment relationship, the employee presumptively

---

17 See Gaius, 149, noting that using asymmetrical allocations of gains and losses are not “against the nature of a partnership,” conceived of in Aristotelian terms. Gaius then goes on to discuss other splits in the contribution of capital and labor in the firm, and notes that the labor of one partner could be so great than no capital contribution is required. Note, however, that the moment the parties depart from the standard pro rata form, the costs of governance are higher.Ironically, these special arrangements are far easier in small businesses than in large firms which tend to have a single class of stock in part to encourage trading on public exchanges.
gets the wages, and only thereafter does the employer receive net profits from the firm. To use the language of capital structures, the employer holds an equity position in the firm, while the employee is in the position of an unsecured creditor.

We thus have profound differences in the payout and control functions, which in fact move hand in hand. Greater control over the business leaves the employer to manage risk and to assume the position of the residual claimant. It is as though the employee is in the position of the holder of a lien and the employer is the holder of the equity. The former gets the higher priority and the letter gets the greater reward for taking the additional risk. There is no doubt of course that the Alchian/Demsetz model is correct to state that the synergistic effects of team production lead people to form firms in the first place. But that explanation does not explain which form of team production is preferable, and why. We need an explanation for the choice of business form in the two-party case, and to set the framework that explains the more complex assignments of risk and reward that can be introduced to tweak the basic arrangement. For example, employees can work on commission or receive bonuses. Some partners do not have an equity interest. Some employees receive commissions or bonuses that depend on the performance of the firm.

So why do some two-person businesses form as partnerships and others as employment relationships? The answer cannot be the division of labor: that could be accommodated in either system. Nor can it lie in the view that the costs of running a firm are cheaper than going to the spot market, for that is also true in both cases. Rather it has to do with variations in competence across individuals. The choice of partners and employees is hardly random. It could easily be that the abler person chooses to pick a lesser party as the employee precisely to avoid the share of control that comes about with the partnership arrangement.

Suppose we know that, of the two persons in question, one has a greater intellectual ability and awareness of general surroundings. The other has lesser abilities and experience: both competence variables line up in the same way. It is not credible to assume that second person will be the employer and the first person will be the employee. The combination of traits indicates that the total gains from cooperation will be greater if the abler person takes both control and an equity position in the revenues of the firm. Both parties know this arrangement, for the ability to recognize that other people are smarter than you are does not depend on your being as smart as they are.

The point has implications for how we think about imperfect information. There are many things that people do not know. These include, however, things that people know that they do not know. The employment relationship thus gives the appropriate debt/equity protection for human capital of those who know that their knowledge is less than that of their boss. Under this arrangement, employee with
less knowledge or ability gets a first call on assets. His protection lies in the fact that
his wages come before his boss’s profits. He both takes less risk and needs less
information to monitor the arrangement in ways that protect his job security and
financial stake. The boss for his part gets the benefit of leverage. By taking the
higher risk he generates the greater return. The fundamental asymmetry in
knowledge and abilities leads to a sensible division of risk, which in turn depends on
a sensible allocation of functions. In more complex structures with multiple
employees these characteristics may lead to different results, as with commissions
and bonuses, but throughout the analysis the same basic factors are likely to be
evident.

The paradigmatic case of the partnership usually starts from a different set of
implicit assumptions about the competence and interests of the parties. In the
simple two party case, the partnership posits two persons with equal ability and
with an equal capacity to take and absorb risks. The business could have strong
specializations in labor, so that one party does the outside marketing and sales and
the other the inside jobs of design and manufacture. In other cases, there could be a
virtual identity of functions, as with two writers who are working as equals on a
joint script. In other situations, the level of specialization may be imperfect, as with
Rogers & Hammerstein, where one party composes music and the other lyrics, but
each side clearly influences the other.

Overall, variations in ability help drive the parties to one form of
organization or another. Yet once the types are cast, the behaviors are not as free
form as the Alchian/Demsetz model suggests because it takes both a strong
justification and a lot of work to renegotiate firm structure. The patterns of control
are highly regular, and neither party wants to jeopardize the benefits of the long‐
term relationship by insisting on an inversion of the basic structure. True, the
prospect of retirement when the partners are not of the same age, and the need to
rework the agreement should the relative competence of the two parties change are
constant risks in business, which in some cases loom so large that dissolution of the
firm becomes the best available option. But whether we look at the snapshot or at
the moving picture, the level of specialization within a firm is driven in part by the
differences in natural endowments and the taste for risk. Of course, there has to be
some form of synergetic gain for parties to enter into this long-term arrangement at
all, but differences in taste and competence help explain the form that the
cooperative venture will take. In sum, the key point remains: the likelihood of
forming an employment relations increases, ceteris paribus, with the differential
abilities of the parties.

There is, of course, the further question of whether certain persons will wish
to be in any kind of arrangement at all. There is a customary division of personality
types between individuals who are lone wolves on the one hand, and those who are
cooperative types on the other. It is likely that these personality types will indicate whether persons are comfortable in working in long-term cooperative arrangements, where high levels of trust are needed (which is true of both the partnership and the employment relationship) or whether they prefer the arm’s length relationships that involve the discrete sale of finished products (no warranties) for a fixed price (no financing) to the cooperative venture. We should expect people to sort themselves out in accordance with their comfort level in different organizations.

**LEGAL IMPLICATIONS**

These basic considerations help clarify key legal issues that surround the employment relationship. As a matter of first principle, the Coasean approach to transaction costs is strongly supportive of a regime of freedom of contract in labor relations between persons on opposite sides of the labor market. These markets are, of course, not perfectly competitive; there are always search and negotiation costs associated with individual deals. Yet at the same time, there are no systematic negative third party effects such as those found with contracts in restraint of trade, e.g. cartelization or horizontal division of territories. And there are no obvious defects of character, which make either fraud or undue influence a substantial risk. The exact terms of these contracts will of course vary widely across industries and firms, in part because of the relative abilities of an employer to evaluate the contribution of an employee, and of the employee to evaluate the conditions associated with various jobs. Put otherwise, labor contracts have many of the same qualities as contracts for the barter of goods; each side is rightly concerned with the package of nonpecuniary benefits that it derives from the other.

These transactional difficulties offer one explanation as to why parties often prefer to enter into at-will contracts. These agreements economize on the cost of separation and they offer each side protection against the opportunistic behavior of the other. It is, moreover, quite common for these contracts to prove stable over long periods of time precisely because the legal arrangements facilitate the cooperation on all sides. The simplicity of the contract at-will, moreover, doubtlessly produces additional gains—relative to other forms of agreement—as the number of employees increase. The arrangement permits all parties to take into account the interactive effects among employees, and permits quick adaption of the work force to the ever changing demands of the market.

More to the point here, the ubiquity of these arrangements suggests that this simple framework works well regardless of the relative competence of the particular individuals involved in the deal. If each party knows something about what to expect from the other side, the agreement will prove stable so long as those
original expectations hold. In addition, the ability to vary wage terms and responsibilities, for example, means that the basic employment-at-will framework is capable of responding to changes in competence levels, positive or negative, over time. Put otherwise, non-at-will agreements have to be highly context specific, and thus are likely to depreciate in value as circumstances change. CEOs often have to negotiate highly particularized contracts because of the huge number of issues—pensions, options, dismissals. But in most settings the premium is on simplicity because the simplest arrangement turns out to be most flexible in the large run of cases in which at-will contracts are used.

In the current legal environment, however, labor contracts are often subject to extensive forms of regulation, including a wide variety of antidiscrimination laws that try to prevent exploitation in standard labor contracts on the ground of race, sex, or national origin. In principle, I believe that competitive markets tend to discipline any firm that makes invidious distinctions among employees on these grounds, because the inefficiencies that they create fall to the firm’s bottom line. The common forms of “discrimination” is therefore much more likely cost-driven, as with the refusal to pay women the same wages and men for dangerous construction jobs on which they have higher risks of injury. The received wisdom on the issue, however starts from the opposite premise that competitive discipline (which is a central part of the Coasean analysis of the firm) does not restrain forms of invidious discrimination, which here cover those market practices that are said to reflect bias and prejudice instead of a concern with profit maximization.

Direct proof of employment discrimination is usually difficult to come by. One explanation is that discrimination is hard to prove because it is actually infrequent. But for many, the assumption that discrimination is a common occurrence leads to the use of various statistical techniques to infer discrimination from the distribution of jobs by category within the work place.

The implicit assumption behind these models is that in a world without employer discrimination we should find a uniform distribution within and across occupations by race, sex, age, and ethnic background. Here the most confident prediction is the most general. But this position is called into question by the above analysis that sees wide variation in aptitudes, tastes and competence by individuals. There is, of course, no reason that these should line up by any of the categories subject to scrutiny under the antidiscrimination laws. But by the same token, once we admit that variations in individual attributes are common, we should not be surprised, and indeed should even expect, differences in preferences and abilities across various categories, as well as within them. The upshot of this orientation is that, in any complex environment, individuals will not be distributed randomly within or across occupations.
The “within” part of the analysis will in general not create huge concerns, but the differences across groups will surely give rise to claims that deal with discrimination, and the attitude we take skewed occupational distributions will determine exactly how we respond to these claims. One possible view is to rely on categorical distinctions in order to impose bans on certain kinds of employment. Sex differences were once thought to justify legal prohibitions on occupational choice. Thus the infamous decision of *Bradwell v. Illinois*\(^\text{18}\) relied on evident sex differences to justify banning women from the practice of law on the ground that the “paramount destiny and mission” of women under divine law marked them for service as wives and mothers.

But surely, even if this view of sex differences could be established a matter of divine truth, it would not justify the ban in question. Why would women or potential employers choose to tempt fate? The legal prohibition would be unneeded given the self-correcting nature of the system: no women would get jobs if none could do the work. In fact, it did not work out that way once the bans were dropped. The ban overlooks variation in competence within groups, and the restriction never makes sense so long as people are able to self-sort. Those men and women who want to tempt fate may well have private information that they can put to great use in any market.

Not only conservative and religious zealots think of these unnecessarily restrictive schemes. In the Progressive era, the “Brandeis brief”—whose purpose was to explain the sociological justification for various limitations on contractual freedom—was introduced to support a selective minimum wage law that applied to women only, in order to protect them from arduous labor.\(^\text{19}\) While the Brandeis brief remains a tool of litigation, under modern notions of equal protection, differential maximum hour laws for women are flatly unconstitutional.\(^\text{20}\) But it would be wrong to assume an antipaternalistic streak in modern judicial decisions. The perceived imbalance by sex in these cases attacks what has been term “class legislation”\(^\text{21}\) because of the explicit use of sex-classifications in creating the legal framework for employment relations. The want of parity could be cured in either of two ways: extend the maximum hour laws to both sexes (at greater social loss), or repeal it altogether. The first of these approaches is clearly more paternalist than the second, where the explicit protection of liberty of contract for all avoids the still greater disruption in labor markets.

\(^{18}\) 83 U.S. 130 (1872).
\(^{19}\) See Muller v. Oregon, 208 U.S. 412 (1908). The upshot of the decision is that all the female workers were fired and replaced with Chinese immigrants. (Friesen and Collins, 1983. For further discussion, see Bernstein, 2010, Ch. 4.)
\(^{20}\) See Reed v. Reed, 404 U.S. 71 (1971).
\(^{21}\) (Gillman, 1995)
It is therefore wholly improper for government to use collective views of class differences offensively to restrict classes of individuals from certain types of employment. By the same token, it is wrong to ignore the level of natural variation by using the law to ban private decisions of employers and employees to sort themselves out by any of these categories. The point is especially important in disparate impact cases, which seek to draw inferences of discrimination without proof of disparate treatment. So long as talents and abilities are differentially distributed—and some surely are—the one sure sign of explicit discrimination is the creation of a wage policy that calls for lock-step increases in occupational categories regardless of race or sex. The assumption of uniformity will lead systematically to the wrong conclusion by finding discrimination where it does not exist.

This basic point is demonstrated quite powerfully for sex differences in occupations in Victor Fuch’s somewhat dated but still instructive book on Women’s Quest for Economic Equality.22 Even within fairly tight occupational classifications (e.g., doctors), the sex differences in the percentage of female dermatologists and neurosurgeons is quite conspicuous, and doubtless reflects at least in part some natural sorting mechanism that, far from being the result of discrimination by employers, instead rests on preferences by employees. That same line of argument was put to effective, if controversial, use in EEOC v. Sears Roebuck & Co.,23 where the company mounted a successful defense against charges of employment discrimination by showing that the women applicants preferred, for a variety of reasons, sales work within stores rather than door-to-door soliciting, even though the latter job paid higher rates under its advantageous commission system.

Scholars working in the Coasean tradition should be sensitive to variations that take place within particular categories of employment. Workers have more fine-grained information than economic analysts, so that the professional categorization is always less rich than the local variations.24 Once we are alert to differences in competence and preferences, we should move much more slowly in imposing regulations on markets whose internal operations we do not understand.

Differentials in competence help to shape the content of the substantive law in two areas that on first view appear to have little in common: securities law and ordinary negligence.

22 (Fuchs, 1989)
23 839 F.2d 302 (7th Cir. 1988)
24 (Heckman, 1998)
Securities Regulation

Modern securities regulation rests on the explicit assumption that the law should make markets accessible to small investors. This approach marks a serious mistake in institutional design, given that individuals are likely to have, by aptitude, training, or experience, highly distinct trading skills. Trying to organize a single voluntary market to accommodate all levels of trading skills is a mistake. It is like trying to organize a public freeway with a bike lane, or a race on the Indianapolis speedway that must also accommodate go-karts. Segregation of markets by abilities has the real advantage of speeding up the level of interaction, because each person knows that it can expect the highest level of performance from others. The situation thus differs from that of ordinary sidewalks, where the expectation of open entry for all is possible precisely because the limitations on speed reduces the need to control the variance in abilities across the population.

Accordingly, social institutions that develop internal norms and background contractual understandings that depend on high levels of sophistication should be binding on all outsiders knowingly entering into any specialized arena. Once again, in certain high performance activities, market separation at a high standard of performance beats market homogenization at a lower standard of performance. That point is instinctively recognized in other settings, such as tournaments organized by age, sex, weight, experience, time, ranking, or whatever. No one would think of putting the slowest runners at the front of a marathon race to neutralize natural differences in ability, which of course marathons are meant to exploit. That same practice sometimes works in job tournaments in which the workers at one level are evaluated to see which subset of them will be promoted.\(^{25}\)

Market Bubbles

Differential levels of ability in securities markets also help to explain a market phenomenon of greater importance—the rapid expansion and complete implosion of market bubbles. There is some real reason to expect that rational actors can foresee the future and should therefore be able to avoid markets that show high levels of instability. Yet the stock market, and real estate bubbles of recent years, can easily be read to convey the opposite message: markets do not

\(^{25}\) (Lazear and Rosen, 1981) The implied tournaments are in part intended to separate works by individual competence. Entrance into the tournament, however, will usually be based on qualifications, which may be established by other tournaments. In any event, the tournament image is not always right. In an expanding organization, many all participants may well be in a position to advance. Strictly speaking all tournaments end up in a single winner.
exhibit these naturally stabilizing characteristics. And there is good reason to believe that instability is a fixture of market behavior, which it is very difficult, if not impossible, to guard against by prudent regulation.

The causal mechanism could easily be this: when markets are stable, the only traders are experienced traders who are not looking for outsized rates of return. But once these markets start to move, less experienced players enter in search of larger gains, and these players do not have the same level of honed instinct as to when a position has been pushed too far. The inexperienced newcomers drive up the market, and the hard question for the experts is how long they want to stay on for the ride. At the initial stages, some will be confident that they can sell to the next generation of buyers independent of any value of the underlying assets. And the higher prices bring in yet other players who know still less. But sooner or later the two pressures reverse. The experienced players head for the exit, each hoping to beat the stampede. The latest players come in with the least experience, but they in turn find that the class of newcomers is exhausted so they have no place to turn, at which point the entire model crumbles.

I see no way to build an explanation about bubbles if we take seriously the assumption that all individual actors satisfy the axioms of rational choice theory. It is only when competence differentials are taken into account that we can get a sensible answer. Moreover, the point seems to have some real empirical foundation. Recent work by Ernan Haruvy and colleagues suggests that as players learn the ropes, the likelihood of asset bubbles reduces, which is exactly as it should be.26 The normative conclusions are much harder to deal with because it is hard to figure out how to keep novel players from coming into open entry markets in which they do not have the skill levels to perform well.

The basic dilemma thus is familiar. The one losing combination is no restrictions on entry coupled with back end protection for those who fall off the edge. The better proposals are to either impose limitations on entry, or allow for voluntary sorting of those who choose to enter.

Automobile Accidents

Segmenting by competence levels is not confined to market settings. This precise issue also arises in determining the applicable standard of care in highway accidents between strangers. In this context, as in others, there are vast levels of difference in individual levels of competence. But it hardly follows that the legal system should calibrate the level of care in accidents to these variations. Indeed, in dealing with accidents between strangers, the standard rule is that both plaintiffs

26 (Haruvy, et. al. 2007)
and defendants have to adhere to the same standard of care, i.e. follow the rules of the road. This approach makes sense is because it is plainly impossible for any driver, no matter what his or her competence, to develop expectations about the conduct of countless other drivers about whom it is impossible to acquire particularized knowledge of competence levels. The effect of this rule is to induce beginners to learn in parking lots under tight supervision, which is as it should be. And in transition—and transitions are always dangerous—new drivers often have markers on the tops of their vehicles so that others can keep at a safe distance, thereby reducing the risk of accident, without changing the rules of the road.

Yet once the individual pairings are known, subjective standards of liability make a good deal more sense. Thus if a given individual agrees to teach a novice how to drive, the teacher knows of the additional risk, which in turn makes it highly likely that the teacher took the risk that the new drive could only meet lower standard of care. Generally speaking in direct social negotiations people do not assume obligations that they know they cannot meet. It follows There is, in a word, no accuracy for any economic model negligence liability that ignores the distinction between the consensual and stranger situations. Rather, the fuller account provides a sensible way in which to sort out transactions in which uniform standards are required from those in which they are not.

**Conclusion**

The purpose of this short paper is to introduce simple notions that help to organize complex spheres of human interaction. The great contribution of Coase is to indicate how transaction costs help explain the familiar forms of behavior that are adopted by various organizations. It would, however, be a mistake to assume that the only drivers of market structure are variations in transaction costs. But whenever we look at labor, financial and product markets, the common sense reaction that people imperfectly rational has some purchase. The point here is not that people are willfully self-destructive, although some are. Rather it is to remind us of this simple point: if everyone except rational choice theorists is aware of the differential capacity and foibles of human beings, it is wise for economists, and even for lawyers, to follow their lead. The former should do it in order to get a better grip of human behavior in a wide range of settings. The latter should do it to be aware of how our limitations should caution us against hasty judgments about matters we don’t understand.

---

28 Holland v. Pitocchelli, 13 N.E.2d 390 (Mass. 1938). For a longer discussion of why different levels of competence imply that there is no single optimal standard of care in other areas of negligence law, see Epstein, 2009.
Ronald Coase was always aware of that problem. He wrote: “if an economist finds something—a business practice of one sort or another—that he does not understand, he looks for a monopoly explanation. And as in this field we are very ignorant, the number of ununderstandable practices tends to be very large, and the reliance on the monopoly explanation frequent.”29 In the final words in a highly instructive paper on price discrimination in competitive markets, Michael Levine added the words: “too frequent, it appears.”30

In many complex industrial models, it is quite unnecessary to relax the model of uniform competence. The forces of selection are too strong. But most of the world does not consist of high-powered exchanges, and so to understand the nature of the firm, and much else beside, we must be aware that rational choice explanations only explain part of the picture. Piecing together the implication of differential competence helps to fill the gap.

---

29 (Coase, 1972)
30 (Levine, 2002)
REFERENCES


EEOC v. Sears Roebuck & Co. 839 F.2d 302 (7th Cir. 1988).


Gaius, Institutes III.


Reed v. Reed, 404 U.S. 71 (1971).


Readers with comments should address them to:

Professor Richard A. Epstein
University of Chicago Law School
1111 East 60th Street
Chicago, IL  60637
repstein@uchicago.edu
Chicago Working Papers in Law and Economics  
(Second Series)

For a listing of papers 1–450 please go to Working Papers at http://www.law.uchicago.edu/Lawecon/index.html

<table>
<thead>
<tr>
<th>Paper Number</th>
<th>Title</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>452.</td>
<td>Richard Epstein, The Case against the Employee Free Choice Act (January 2009)</td>
<td></td>
</tr>
<tr>
<td>453.</td>
<td>Adam B. Cox, Immigration Law’s Organizing Principles (February 2009)</td>
<td></td>
</tr>
<tr>
<td>454.</td>
<td>Philip J. Cook, Jens Ludwig, and Adam M. Samaha, Gun Control after Heller: Threats and Sideshows from a Social Welfare Perspective (February 2009)</td>
<td></td>
</tr>
<tr>
<td>455.</td>
<td>Lior Jacob Strahilevitz, The Right to Abandon (February 2009)</td>
<td></td>
</tr>
<tr>
<td>457.</td>
<td>Lee Anne Fennell, Commons, Anticommons, Semicommons (February 2009)</td>
<td></td>
</tr>
<tr>
<td>464.</td>
<td>Anupam Chander, Corporate Law’s Distributive Design (June 2009)</td>
<td></td>
</tr>
<tr>
<td>465.</td>
<td>Anupam Chander, Trade 2.0 (June 2009)</td>
<td></td>
</tr>
<tr>
<td>467.</td>
<td>Eric A Posner, Kathryn Spier, and Adrian Vermeule, Divide and Conquer (June 2009)</td>
<td></td>
</tr>
<tr>
<td>468.</td>
<td>John Bronstein, Christopher J. Buccafusco, and Jonathan S. Masur, Welfare as Happiness (June 2009)</td>
<td></td>
</tr>
<tr>
<td>469.</td>
<td>Richard A. Epstein and Amanda M. Rose, The Regulation of Sovereign Wealth Funds: The Virtues of Going Slow (June 2009)</td>
<td></td>
</tr>
<tr>
<td>470.</td>
<td>Douglas G. Baird and Robert K. Rasmussen, Anti-Bankruptcy (June 2009)</td>
<td></td>
</tr>
<tr>
<td>472.</td>
<td>Bernard E. Harcourt, Neoliberal Penalty: A Brief Genealogy (June 2009)</td>
<td></td>
</tr>
<tr>
<td>473.</td>
<td>Lee Anne Fennell, Willpower and Legal Policy (June 2009)</td>
<td></td>
</tr>
<tr>
<td>475.</td>
<td>Randal C. Picker, Online Advertising, Identity and Privacy (June 2009)</td>
<td></td>
</tr>
<tr>
<td>476.</td>
<td>M. Todd Henderson, Credit Derivatives Are Not “Insurance” (July 2009)</td>
<td></td>
</tr>
<tr>
<td>477.</td>
<td>Lee Anne Fennell and Julie Roin, Controlling Residential Stakes (July 2009)</td>
<td></td>
</tr>
<tr>
<td>481.</td>
<td>Lee Anne Fennell, The Unbounded Home, Property Values beyond Property Lines (August 2009)</td>
<td></td>
</tr>
<tr>
<td>484.</td>
<td>Omri Ben-Shahar, One-Way Contracts: Consumer Protection without Law (October 2009)</td>
<td></td>
</tr>
<tr>
<td>485.</td>
<td>Ariel Porat, Expanding Liability for Negligence Per Se (October 2009)</td>
<td></td>
</tr>
<tr>
<td>486.</td>
<td>Ariel Porat and Alex Stein, Liability for Future Harm (October 2009)</td>
<td></td>
</tr>
<tr>
<td>487.</td>
<td>Anup Malani and Ramanan Laxminrayan, Incentives for Surveillance of Infectious Disease Outbreaks (October 2009)</td>
<td></td>
</tr>
</tbody>
</table>
488. Anup Malani, Oliver Bembom and Mark van der Laan, Accounting for Differences among Patients in the FDA Approval Process (October 2009)
489. David Gilo and Ariel Porat, Viewing Unconscionability through a Market Lens (October 2009)
491. M. Todd Henderson, Justifying Jones (November 2009)
497. Randal C. Picker, Easterbrook on Copyright (November 2009)
498. Omri Ben-Shahar, Pre-Closing Liability (November 2009)
500. Saul Levmore, Ambiguous Statutes (November 2009)
501. Saul Levmore, Interest Groups and the Problem with Incrementalism (November 2009)
503. Nuno Garoupa and Tom Ginsburg, Reputation, Information and the Organization of the Judiciary (December 2009)
506. Richard A. Epstein, Impermissible Ratemaking in Health-Insurance Reform: Why the Reid Bill is Unconstitutional (December 2009)
512. Omri Ben-Shahar and Anu Bradford, The Economics of Climate Enforcement (February 2010)