COMPETING NORMS AND SOCIAL EVOLUTION: IS THE FITTEST NORM EFFICIENT?

PAUL G. MAHONEY* & CHRIS W. SANCHIRICO†

An influential theme in recent legal scholarship is that law is not as important as it appears. Social control, many scholars have noted, is often achieved through social norms—informal, decentralised systems of consensus and cooperation—rather than through law. This literature also displays a guarded optimism that social evolutionary processes will tend to favour the adoption of efficient norms. Using concepts from evolutionary game theory, we demonstrate that efficient norms will prevail only in certain settings and not in others: survival of the fittest does not imply survival of the efficient. In particular, we show that in many games of interest to legal scholars—games describing fundamental interactions in property, trust, and contract—evolutionary forces lead away from efficiency. We also describe how law rights this trend.

INTRODUCTION

Economists are paying increasing attention to the role of institutions—notably legal institutions—in the operation of economic systems. Both economic historians and financial economists contend that legal rules protecting investment are essential for markets to flourish.1

At the same time, an influential theme in legal and political science scholarship is that law is not as important as it appears. Social

† Brokaw Professor of Corporate Law and Albert C. BeVier Research Professor, University of Virginia School of Law.
‡ Associate Professor, University of Virginia School of Law.

We thank Jenna Bednar, George Cohen, Robert Ellickson, John Harrison, Jack Knight, Saul Levmore, Julie Mahoney, Francesco Parisi, Eric Posner, John Setear, Jonathan Wiener, and participants in the Norms and Corporate Law Symposium at the University of Pennsylvania, workshops at the University of Michigan Law School and the University of Texas Law School, and a panel at the American Law & Economics Association, 2000 annual meeting. Rachel Brewster provided valuable research assistance.

1 See DOUGLAS C. NORTH, STRUCTURE AND CHANGE IN ECONOMIC HISTORY 6 (1981) (“The security of property rights has been a critical determinant of the rate of saving and capital formation.”); Rafael La Porta et al., Law and Finance, 106 J. POL. ECON. 1113 (1998).
control, many scholars note, is often achieved through social norms—informal, decentralized systems of consensus and cooperation—rather than through law. Many discussions of norms draw on a separate literature on cooperation in game theoretic settings, particularly in repeated games. Using formal or informal models of norms as strategies in repeated games, scholars, such as Robert Ellickson, express a tentative optimism that socially desirable, or efficient, practices can arise from the interactions of decentralized, utility-maximizing individuals. Moreover, Randal Picker and Robert Axelrod argue not just that efficient norms are the possible outcomes of decentralized interactions, but that such norms have, by virtue of their efficiency, a natural advantage in the social evolutionary processes that determine norm adoption.

These seemingly contradictory positions may be reconciled so long as each is kept within its proper domain. Using concepts from evolutionary game theory, we demonstrate that efficient strategies will prevail in evolutionary settings for some games, but not for others. In particular, we show that many games of interest to legal scholars—games that are analogous to fundamental interactions in property, tort, and contract—are precisely those in which evolutionary forces lead away from efficiency, and institutions are required to right the trend.

At the heart of the distinction between evolutionary “fitness” and economic efficiency lies the concept of mismatch risk. Whether a norm is efficient depends entirely on how well it does when matched against itself. Norm A is Pareto superior to norm B if everyone has higher payoffs when norm A is generally adopted. But whether a

---


norm is selected by evolutionary processes depends also on how well it
does when mismatched with other norms. On the evolutionary
battlefield, the successful norm, like the successful weapon, must also
fare comparatively well against those not similarly armed. A norm
that is Pareto optimal but does relatively poorly against disparate
norms is said to have mismatch risk. When a norm’s mismatch risk is
larger than its efficiency advantage, evolution will tend to select
against it.

We argue that mismatch risk is a plausible metaphor for many
important real-world problems that arise because efficient practices
require investment and inefficient practices may involve appropriating
that investment or destroying much of its value. Thus, while the
investment-oriented norm may be efficient, it is likely to be the
relative loser when the other side reneges or appropriates. The most
basic legal rules—those providing protection for property rights,
enaforcing contracts, and requiring reasonable care in the conduct of
risky activities—can be seen as responses to just such situations.

When mismatch risk is great in relation to the potential gains
from playing efficient strategies, completely decentralized repeat play
is unlikely to produce efficient norms. To assure efficiency, we must
introduce an external party that can alter payoffs. We note that in
well-functioning societies, law acts as this third party. It redistributes
part of the defecting party’s payoff to the cooperating party, thus
mitigating mismatch risk and improving the evolutionary fitness of
efficient strategies. Therefore, while norms scholars are correct to
point out that law is not the only, or even the primary, means of social
control, economists are also right to point out that law and efficiency
are usually found together.

The rest of this paper is organized as follows. In Part I, we
summarize and connect the relevant strands of the legal literature on
norms, emphasizing the optimistic arguments about norm evolution.
Part II presents a nontechnical analysis of the parallel literature on
evolutionary game theory. Here, our emphasis is on the surprisingly
general agreement within that literature that efficiency and fitness are
distinct concepts. We also discuss that relationship between fitness
and efficiency is affected when agents can choose with whom they
interact. In Part III, we provide several elementary legal examples in
which mismatch risk plays an important role. Part IV explores the
consequent necessity of centralized institutions in guiding
evolutionary processes toward efficient outcomes.
I. NORMS: COOPERATION WITHOUT LAW

Norms, as the term has come to be used in legal scholarship, are rules of conduct that constrain self-interested behavior and that are adopted and enforced in an informal, decentralized setting.8 They arise in the course of individuals' repeated interactions, which sometimes produce a consensus about desirable and undesirable conduct.9 Enforcement of the resulting rules takes place outside the legal system through mechanisms such as gossip or ostracism. If a norm becomes sufficiently strong, individuals may internalize, or self-enforce, the norm by feeling guilt when they contemplate its violation.10 The norms literature has produced a detailed taxonomy of sanctions, including "first-party" sanctions such as feelings of guilt or shame, "second-party" sanctions such as retaliation by a victim, and "third-party" sanctions such as social disapproval.11

A central issue in the norms literature is whether norms will generally be efficient. We would characterize the dominant theme among scholars who have studied the issue as guarded optimism, although a few pessimistic notes have been sounded as well.12 We briefly survey the case for the efficiency of norms.

A. Norms as Repeated Game Equilibria

The leading theoretical approach to the efficiency question (indeed, to the entire law and economics literature on norms) is the theory of repeated games. Norms are modeled as equilibrium strategy

8 Similar definitions are employed elsewhere in the social science literature. See, e.g., Jean Elsasser & Jack Knight, Changing Social Norms: Common Property, Braggadocio, and Gang Enquiry, 58 CURRENT ANTHROPOLOGY 1, 2 (1997) ("[S]ocial norms are informal rules that structure behavior in ways that allow individuals to gain the benefits of collective action.").
9 See Robert D. Cooter, Structural Adjustment and the New Law Merchant: A Model of Decentralized Law, 14 INT'L REV. L. & ECON. 215, 224 (1994) ("When a public consensus forms that people ought to cooperate, and when enough people internalize the obligation to punish non-cooperators effectively, a social norm exists in the community.").
10 Gary S. Becker, ACCOUNTING FOR TASTES 159-60 (1996).
12 E.g., Picker, Simple Games, supra note 4; Ellikson, Evolution, supra note 3.
13 See Jody S. Kraus, Legal Design and the Evolution of Commercial Norms, 30 J. LEGAL STUD. 577, 585 (1997) (arguing that norms may be "significantly suboptimal"); Eric A. Posner, Law, Economics, and Inefficient Norms, 144 U. PA. L. REV. 1697, 1743 (1996) ("[E]ven close-knit groups are likely to produce inefficient norms . . . .").
choices is a particular repeated game. Depending on the situation, other authors would narrow the definition by limiting it to cooperative strategies or to collectively enforced strategies. These refinements are unnecessary for our purpose, which is to describe the relative evolutionary fitness of cooperative and uncooperative strategies.

Thus, for the sake of simplicity, we use a broad definition.

To provide some real-world motivation, we can imagine a pair of neighbors. An attractive tree straddles their property line. The tree requires some care—pruning, mulching, and so on. That care generates a positive externality. Part of the benefits of each individual's effort accrues to her neighbor; and while the joint benefits from individual effort exceed private costs, private benefits do not. To gain tractability at the expense of some verisimilitude, we limit each neighbor to two possible actions chosen in isolation from the other—pay $50 for care or pay nothing. When only one neighbor pays $50 for tree care, each neighbor individually enjoys a $40 benefit, while if both pay $50, each enjoys an $80 benefit. The resulting payoffs are shown in Table 1. Throughout, we will show only the row player's payoffs in the matrix representations. All of the games we will use are symmetrical in the sense that the row and column players are interchangeable. We indicate pure strategy Nash equilibria by shading.

<table>
<thead>
<tr>
<th></th>
<th>Pay</th>
<th>Don't Pay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pay</td>
<td>30</td>
<td>-10</td>
</tr>
<tr>
<td>Don't Pay</td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Tree Maintenance as a Prisoner's Dilemma

The game is a prisoner's dilemma and the dominant strategy for

---

11 For examples of this approach, see ELLICKSON, ORDER, supra note 3, and COOPER, supra note 6.
12 See ENSMINGER & SNIGHT, supra note 5, at 2 ("[Social] norms are informal rules that structure behavior in ways that allow individuals to gain the benefits of collective action.").
13 See Bendor & Swiat, supra note 4, at 322 (discussing strategies in which "I can punish 2 for 2's conduct toward third parties").
each neighbor is to pay nothing. The outcome is Pareto inefficient, as each neighbor could be made better off if each paid $50. Although the situation is bleak if the players meet each other only in a one-shot interaction, the well-known folk theorem of repeated games suggests that cooperation in caring for the tree is a possible outcome if the neighbors play this game repeatedly (more generally, are engaged in an ongoing relationship). Specifically, if each individual can condition current actions on her neighbor’s past behavior, and cares not just about current payoffs but, to a sufficient degree, about future outcomes, then cooperation can be maintained as a Nash equilibrium of the infinitely (or indefinitely) repeated game.14

One salient and compelling mode of cooperation is the “tit-for-tat” strategy. According to this strategy, each neighbor pays to maintain the tree in the first year and from thereon imitates her neighbor’s prior-year behavior. Provided the neighbors care enough about future payoffs, tit-for-tat is an equilibrium of the repeated game. It is also clear that if both neighbors play this strategy, they will enjoy the higher payoffs of perpetual cooperation.

We should note here that we are side-stepping yet another definitional issue that has plagued the norms literature.15 Some scholars use the term “norm” to refer to a common expectation and practice regarding behavior in a particular relationship setting.16 Others use it to describe the system of social sanctions that are applied when an individual defects from this common expectation and practice.17 Fortunately, the game theoretic notion of a strategy encompasses both of these concepts. A strategy for an individual specifies how that individual will behave in each contingency as a relationship unfolds over time. Interpreted in light of the individual’s expectations of others’ strategies, a strategy includes the individual’s behavior both when she sees the relationship proceeding on course

---

14 For a formal review, including a discussion of subgame perfection and renegotiation-proofness of repeated game equilibria, see DREW PUDENBERG & JEAN THOILL. GAME THEORY (1991).

15 Robert Scott has aptly noted that norms literature “suffers from conceptual pluralism and terminological disarray.” Robert E. Scott, The Limits of Behavioral Theory of Law and Social Norms, 86 WIS. L. REV. 1605, 1607 (2000). We aspire not to make the problem worse.


17 See Ellipsoid, Evolution, supra note 3, at 8 (“assuming that norms as ‘social sanctions can range from doing nothing, to making a comment or facial gesture, to administering a weightier response such as a gift or punch to the jaw’.”).
and when she reacts to another individual’s defection. Thus, given the relationship’s behavioral history, an individual’s strategy describes what we might call her “primary behavior,” what she will do based on her current expectations of others’ choices, as well as her “secondary behavior,” what she plans to do if others disappoint these expectations. In the tree-maintenance game just described, the tit-for-tat strategy specifies that the individual begins the game by cooperating, evidenced here by paying. Interpreted in light of the individual’s expectation that others also follow tit-for-tat, the strategy also specifies what the individual plans to do if others disappoint her expectations and fail to cooperate. In this way, the tit-for-tat strategy specifies both primary and secondary conduct at the “null history” of play; that is, at the commencement of the relationship. The same can be said with respect to any past history of play.

When the subject of our evolutionary analysis is a relationship that explicitly unfolds over time—such as the tree maintenance game—we use the term “norm” to refer to an equilibrium strategy profile: a list of strategies, one for each player, such that each player’s strategy matches every other player’s expectations thereof and is the best response to these expectations. Thus, our notion of a strategy encompasses both primary and secondary behavior.

It is also worth noting that in many settings the range of secondary behavior may differ in kind from the range of primary behavior—as when someone reacts to her neighbor’s failure to cooperate in tree maintenance by gossiping or reducing her own cooperation in other neighborly interactions, such as taking in the newspaper when the neighbor is on vacation or repairing a common boundary fence. In all of the games we analyze, we model all secondary behavior with the identical game form as the primary behavior. In the tree-maintenance game, for example, we assume that all interactions between the neighbors can be modeled as identical prisoner’s dilemma games. Thus, every round of play is another “tree-maintenance” game even if the actual actions are something else (such as gossiping or refusing to repair a fence).

We also assume that each player plays the same strategy against all opponents. Thus, if Neighbor A plays a repeated prisoner’s dilemma game with Neighbors B, C, and D, Neighbor A chooses a single strategy to play vis-à-vis all 3 neighbors. As a consequence, there is no distinction between “second-party” sanctioning (retaliation by the person harmed directly by a defection in the prior round) and “third-party” sanctioning (sanctions by other players who were not directly
affected by the defection).

These simplifications make it considerably easier to describe the repeated games and their equilibria. They are also consistent with the tenor of the informal repeated game frameworks typically used in the norms literature, as well as Axelrod’s and Picker’s more formal analyses. Most importantly, however, they enable us to focus on repeated game dynamics applicable to a rich array of real-world settings with a minimum of definitional baggage.

As we have shown, the traditional analysis of the repeated prisoner’s dilemma provides reason for optimism that norms (equilibria) involving cooperation can exist. Yet this optimism does not go so far as to imply that people will generally choose the one strategy, from a set of possible strategies, that provides the highest average payoffs. An individual will choose the strategy that is best for her against the current strategy composition of the group, not necessarily the strategy that would be best for all were all to adopt it. In particular, although tit-for-tat is an equilibrium of the repeated game, so is the strategy “always defect” in which the neighbor never pays to care for the tree no matter what the history of play. The theory of repeated games by itself provides us with no reason why one outcome and not the other will obtain. Consequently, if it is to be a source of optimism regarding the adoption of efficient norms, a repeated game analysis must be coupled with some notion of how norms fare in competition with one another.

B. Evolutionary Approaches

1. The Evolution of Cooperation in Axelrod

Axelrod’s seminal research provides cause for optimism about the emergence of cooperative (efficient) strategies.16 In his classic round robin computer tournament, tit-for-tat consistently outperformed all other submitted strategies (in terms of average payoffs over all pairwise encounters). Generalizing from this experimental result, Axelrod concludes that tit-for-tat’s internal structure makes it the likely outcome of evolutionary processes in which successful strategies are imitated and unsuccessful strategies abandoned. More recently, Bendor and Swistak confirmed that Axelrod’s central assertion regarding tit-for-tat’s evolutionary advantage does indeed submit to

---

16 AXELROD, supra note 4; see Bendor & Swistak, supra note 4, at 281 (formalizing and clarifying Axelrod’s results).
formal mathematical proof.\textsuperscript{10}

Axelrod's results have great intuitive appeal because defection reflects, albeit abstractly, a type of behavior familiar to most of us. Not surprisingly, the norms literature looks to Axelrod's results as a hopeful sign that norms of cooperation will emerge in small communities of repeat players.\textsuperscript{11} The community may develop a consensus that a particular behavior (such as shoveling the sidewalk outside one's house when it snows) is desirable. Failure to comply with the consensus can prompt retaliation in the form of failure to cooperate in other endeavors.

2. The Evolution of Cooperation in Picker's Model

In an important recent paper, Picker corroborates Axelrod's optimism from a different starting point.\textsuperscript{12} Picker's models begin with a 2 x 2 coordination game, that is, one in which each of the two strategies is a best response to itself. Picker's basic schematic is shown in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Norm A</th>
<th>Norm B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norm A</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Norm B</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Norm Selection as a Coordination Game ($a, b > 0$)

The game has two Nash equilibria where both players select Norm A or both select Norm B. The simplest example would be driving on a divided road. Payoffs are highest when everyone drives on the left or

\textsuperscript{10} Axelrod's main intuition, that retaliatory strategies of conditional cooperation are somehow advantaged, proves correct in one specific and significant sense. Under a standard evolutionary dynamic these strategies require the minimal frequency to stabilize. Hence, they support the most robust evolutionary equilibrium: the easiest to reach and retain. Moreover, the less efficient a strategy, the larger is its minimal stabilizing frequency. Hobbesian strategies of pure defection are the least robust. Bendor & Swinkels, supra note 4, at 290; see also id. at 299-301.

\textsuperscript{11} See Elliston, Order, supra note 3, at 910.

\textsuperscript{12} See Picker, Simple Games, supra note 4, at 1220-29.
on the right, and either is a Nash equilibrium.

When selecting a rule to govern driving on a divided road, neither equilibrium is better than the other. In terms of the model, \( a = b \).

That need not be true generally. The interesting question is, when \( a > b \) or \( b > a \), will players select the efficient (higher payoff) strategy? \(^2\)

Picker's evolutionary model posits that players sometimes randomly switch strategies, or mutate. He begins with a population of individuals distributed one to a cell on a torus-shaped grid. In every round of the repeated game, each player plays the game with each neighbor in the 8 cells adjacent to its own. After observing the resulting payoffs, he imitates "successful" strategies (which can be defined in different ways) in future rounds.

The results vary, of course, depending on how much better the efficient norm is than the inefficient norm, how many players initially play each strategy, the rate of mutation, and the rules players use to determine whether to switch strategies deliberately. Overall, however, Picker shows that over a reasonably broad range of parameters, evolution selects the efficient strategy—over time, more players deliberately switch from the inefficient to the efficient strategy than vice versa. \(^2\)

3. Connecting Axelrod and Picker

At first glance, Picker's analysis seems unrelated to Axelrod's—indeed, unrelated to the rest of the norms literature that focuses on repeated prisoner's dilemmas. Picker's underlying game form is a coordination game, not a prisoner's dilemma, and his players do not choose their strategies in the current round based on opponents' actions in prior rounds, but rather based on what strategy will provide the highest payoff given an assumed neighborhood strategy composition. But Picker's analysis extends readily to the case of a repeated prisoner's dilemma game and when so extended goes to the

---

\(^2\) Picker analyzes the same coordination game in Picker, Neighborhoods, supra note 4. In this more recent paper, however, the grid is not entirely full (thus impeding the chain reaction-like transmission of successful strategies across interlocking neighborhoods). Further, players who do not meet their (exceptional) aspiration levels in their current location can relocate to an empty spot on the grid, one that will be randomly chosen for them. In evaluating the potential for the dissipation of successful coordination, Picker emphasizes the importance of either or both low aspiration levels (and thus frequent relocation) and dense networks. The model differs from the recent work on mobility in evolutionary game theory as discussed infra in Part II.C. In that literature, unlike Picker's model, agents may choose where to locate when they re-optimize.
heart of the matter: namely, whether efficient repeated game strategies will be adopted. In this respect, Picker’s analysis builds upon Axelrod’s explanation of how and why repeated game cooperation will likely evolve.

To see how the repeated prisoner’s dilemma game studied extensively in the norms literature can be collapsed into the sort of coordination game Picker studies,26 consider the particular prisoner’s dilemma payoffs used in Axelrod’s tournament:

<table>
<thead>
<tr>
<th></th>
<th>Cooperate</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Defect</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Axelrod’s Prisoner’s Dilemma

In analyzing the indefinitely repeated versions of this game, we will assume for simplicity that the players are required to choose between two repeated strategies—“tit-for-tat” or “always defect.” (The same principles apply if all strategies are included.) In Axelrod’s second tournament, the duration of each pair-wise encounter was randomly determined.27 In particular, the chance of going yet another round with one’s current opponent was always \( w = 99.684\% \). This meant that (risk neutral) players discounted future payoffs with discount factor \( 0.99684 \) (that is, a discount rate of \( r = 0.347\% \)). The expected payoffs from each repeated encounter can then be easily calculated.28 Two tit-for-tatters would each receive 3 per period, for an expected payoff of \( 3/(1-w) = 387 \). Similarly, two always defectors would each receive \( 1/(1-w) = 289 \). And if a tit-for-tatter met an always defector, the tit-for-tatter would receive \( 0 + w(1/(1-w)) = 288 \), and the always defector would receive \( 5 + w(1/(1-w)) = 293 \). These payoffs are

---

26 For the general idea that extensive form games have a strategic (normal) form representation, see, for example, FUDENBERG & TIBORI, supra note 14, at 85-87.

27 In Axelrod’s first tournament, the prisoner’s dilemma was played for 200 repetitions for each pair-wise encounter with no discounting of future payoffs. The finite nature of the game produced complicating endgame effects. The second tournament corrected for these complicating effects. See AXELROD, supra note 4.

28 Recall that if \( 0 < a < 1 \), then \( 1 + a + a^2 + a^3 + \ldots = 1/(1-a) \). Id. at 15.
shown in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>Tit-for-Tat</th>
<th>Always Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tit-for-Tat</td>
<td></td>
<td>288</td>
</tr>
<tr>
<td>Always Defect</td>
<td>223</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Axelrod's Second Tournament Simplified: Discounting at \( w = 0.99654 \)

Tit-for-tat is the best strategic response to itself, as is always defect. Further, tit-for-tat Pareto dominates always defect. The "collapsed" repeated prisoner's dilemma in Table 4, therefore, is qualitatively similar to Picker's norm coordination games. And, in fact, if we were to apply Picker's model to Table 4, we would obtain a similarly optimistic result regarding the evolutionary dominance of the efficient norm. We would also corroborate, in a more fully developed model, the preliminary evolutionary analysis that Axelrod supplies after reporting tournament results.

Picker's norm coordination game and Axelrod's collapsed repeated prisoner's dilemma share another notable characteristic aside from the existence of two Pareto-ranked equilibria. In both games, the efficient strategy entails no mismatch risk. When a player adopting the efficient norm meets a player who adopts the inefficient norm, both players suffer roughly equally for the mismatch. This is manifest in the fact that the off-diagonals payoffs in both tables—the payoffs in the southwest and northeast cells of the respective matrices—are roughly the same. In Picker's norm coordination game they are both zero. In Axelrod's collapsed repeated game they are within 5 units of each other, in a game whose payoffs are measured in the hundreds.

As we shall see, this common characteristic goes a long way toward explaining Picker's and Axelrod's common optimism about the evolution of efficient norms. The structure of the off-diagonal payoffs

\[ \text{Indeed, both strategies are equilibria in the repeated game when all possible strategies are considered as alternatives.} \]
is, in general, a key determinant of results in evolutionary game theory models, to which we now turn.

II. EVOLUTIONARY GAME THEORY AND MISMATCH RISK

The substantial and growing literature on learning and evolution in games is perhaps the most significant development in game theory over the last decade. Like scholarship on social norms, an important segment of this literature asks which among a set of competing norms—or Nash equilibria—evolutionary processes will favor. In particular, it asks whether evolutionary forces will select the most efficient norms. On this issue the game theory literature displays an unusual uniformity. The consensus is that survival of the fittest is not the same as survival of the efficient. In this Part, we provide a nontechnical account of this literature, focusing on the genesis of and explanation for this inefficiency result.

A. From Game Theory to Biology

Evolutionary game theory grew out of an intellectual volley between theoretical biology and game theory. The history begins with an attempt by theoretical biologists John Maynard Smith and George Price to understand ritual combat within species. Unsatisfied with explanations provided by conventional evolutionary theory, Maynard

---


FUDENBERG & LEVINE, supra note 27, at 186, provides this assessment: "There is one important restriction on 2 x 2 games that deserves emphasis, since it obtains in many (but not all) such models. This is the selection of the risk-dominant equilibrium (the equilibrium consisting of the best action if one's opponent is equally likely to play either action) as the unique long-run steady state. Of particular importance is the connection (and lack of connection) of risk-dominance to Pareto efficiency. In pure coordination games the two concepts are the same. However, in general games, risk-dominant equilibria may fail to be Pareto efficient. The conclusion from the study of stochastic adjustment models is that learning procedures tend to select equilibria that are relatively robust to mutations (risk-dominant equilibria), and this is a different criterion than Pareto efficiency."

Smith and Price borrowed ideas from game theory, modifying the interpretation of key concepts to fit the contours of animal evolution. A decade or so later, many game theorists became dissatisfied with the foundational assumptions underlying Nash equilibrium, in particular the requisite assumption that players somehow know what their opponents are going to do. In seeking explanations for the genesis of this common misunderstanding of strategic intent, game theorists turned to the concepts of learning and evolution, thus borrowing back Maynard Smith and Price's earlier adaptation of the game form to animal evolution. The result is the field of evolutionary game theory. For our purposes, Maynard Smith and Price's two main contributions were an evolutionary reinterpretation of the basic game construct and a translation of the fundamental phenomena of game theoretic interaction to the evolutionary setting. After explaining

---

30 This dissatisfaction first appeared in the definition of what was computably a new solution concept: "rationalizability." See B. Douglas Bernheim, Rationalizable Strategic Behavior, 52 ECONOMETRICA 1007, 1007 (1983) (criticizing the three-conventional view that Nash equilibrium is implied by common knowledge of rationality among the players); David G. Pearce, Rationalizability Strategic Behavior and the Problem of Perfection, 52 ECONOMETRICA 1029, 1030 (1984) (simultaneously and independently reaching the same conclusion regarding Nash equilibrium as Bernheim in Rationalizable Strategic Behavior and adding a richer analysis of extensive form games). The same dissatisfaction contributed to the development of a literature on learning in game theoretic settings. See Chris N. Sanchirico, A Probabilistic Model of Learning in Games, 64 ECONOMETRICA 1975 (1996) (presenting a theory of how repeated interaction may lead to the common understanding of strategic intent necessary to support Nash equilibrium play).

31 Later work has all but unseated the particular solution concept, the evolutionarily stable strategy ("ESS"), that Maynard Smith and Price offered as a prediction for how the reinterpreted game would proceed. Though the ESSs of a game are meant to be the possible settling points of an evolutionary process, Maynard Smith and Price did not formally define the dynamics of this process. Later research did posit and characterize a dynamic process, the "replicator dynamic," that appeared to be consistent with what Maynard Smith and Price had in mind. For a D. Taylor & Leo B. Jonker, Evolutionary Stable Strategies and Game Dynamics, 40 MATHEMATICAL BIOSCIENCES 145 (1978). But these researchers found that the ESS concept was too restrictive with respect to this dynamic: although all ESSs are resting points of the replicator dynamic, not all resting point of the replicator dynamic are ESSs. Weirbull, supra note 27, contains an exceedingly clear description of the literature up to this point. Taylor and Jonker's solution concept was itself limited by the fact that it did not consider ongoing mutation—the replicator dynamic is a deterministic process, whereas the Maynard Smith and Price model seems to envision ongoing uncertain mutations. More recent work accounts for ongoing mutation. See Dean Foster & Peyton Young, Stochastic Evolutionary Game Dynamics, 28 THEORETICAL POPULATION BIOLOGY 219 (1990); Mashihiro Randorv et al., Learning, Mutation, and Long Run Equilibria in Games, 61 ECONOMETRICA 29, 50 (1993).
these phenomena in the traditional setting, we show how they manifest in the biological model.

1. Payoff Interdependence in the Canonical Game Setting

In a game theoretic setting, each player’s strategy choice influences the other’s payoff. This payoff interdependence makes the setting a “game” rather than two adjacent individual decision problems. Interdependence leads to a departure from the usual invisible hand result—the aggregation of individually maximizing behavior does not necessarily lead to a socially optimal outcome.

We observe this “pessimistic” result most readily in the one-shot prisoner’s dilemma. In order to consider the choice among competing norms in an evolutionary setting, however, we must start with a game that has more than one Nash equilibrium. We accordingly consider the classic Stag Hunt game, a two-player game with two pure strategy equilibria or norms (we ignore mixed strategy equilibria for the time being).

Table 5 shows typical payoffs for the Stag Hunt in the usual matrix representation. The mnemonic parable for this game is as follows: two hunters in search of game hide separately in the brush near a watering hole. If both cooperate, they will catch a stag, which they will share for dinner. Alternatively, each may opt out of the cooperative venture by hunting hare on her own. If one hunter opts out and the other continues to pursue stag, the one who opts out catches a hare while the solitary stag hunter fails to catch her prey and goes hungry. Both hunters prefer to share a stag rather than eat hare alone, so that if each can rely on the other’s cooperation, each will choose to hunt the stag. If the other thinks her partner is likely to opt for hare, however, her best response is to hunt for hare.

The pure strategy Nash equilibria are both hunt for stag or both hunt for hare. Traditional solution concepts do not give us any basis for determining which equilibrium will obtain. This is true even though one strategy is Pareto superior (because both hunters prefer half a stag to one hare).
<table>
<thead>
<tr>
<th></th>
<th>Hunt Stag</th>
<th>Hunt Hare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hunt Stag</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Hunt Hare</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Stag Hunt Game

2. Maynard Smith and Price’s Evolutionary Reinterpretation

In modifying the game to fit the biological context, Maynard Smith and Price view strategies not as choices made by individuals—individual animals, that is—but as hard-wired, genetically inherited modes of behavior. In a large population, players meet each other one-on-one in random fashion to play the subject game according to how they are genetically programmed. The more prevalent a given strategy in the population, the more likely each individual’s strategy will be tested against it. Payoffs in the reinterpreted game are not money or utility, but rather reproductive opportunity or “fitness.” An individual whose actions consistently garner larger payoffs has more offspring who inherit the parent’s strategy. Accordingly, if one action consistently achieves larger payoffs than another, it grows at a faster rate and gradually comes to dominate the population. Maynard Smith and Price also introduced the idea of “mutation”: the rare and random adoption of certain actions regardless of their reproductive fitness is the current population. Thus, despite the absence of consciously interactive decisionmaking, the Maynard Smith and Price model is fundamentally strategic, or at least interactive.

B. From Biology Back to Game Theory

For our purposes, evolutionary game theory makes two important contributions to the Maynard Smith and Price analysis. First, it alters the biological game to fit human interaction (of a sort, as we shall see). Second, it explicitly models the evolutionary dynamics, which are only implicit in Maynard Smith and Price (and in Axelrod’s

83 See id at 16.
evolutionary interpretation of the emergence of tit-for-tat). This second advance is important, because considering the actual stochastic dynamics sharpens the distinction between fitness and efficiency.

1. The Social Re-reinterpretation of the Biological Model

The evolutionary components of a biological model are genetic inheritance and mutation. Literally interpreted, these are unlikely to be applicable to human interaction on a time scale of interest to most game theorists and legal scholars. The model can nevertheless be adapted to describe the evolution of human conventions among populations of boundedly rational individuals in "social science time." For example, we can imagine that individuals are randomly paired to interact according to a fixed game form. But payoffs are again payoffs in the sense of utilities or benefits, rather than opportunities for reproduction. Further, players are no longer hard-wired to play inherited strategies, rather, a naive form of occasional optimization takes over the role played by genetic replication. In

---

1 Axelrod was among the first to propose this reinterpretation. See Axelrod, supra note 4; see also Kenneth G. Binmore et al., Musical Chairs: Modeling Nasty Evolution, 11 GAMES & ECON. BEHAV. 1 (1990); Foster & Young, supra note 51; D. Fudenberg & C. Harris, Evolutionary Dynamics with Aggregate Shocks, 57 J. ECON. THEORY 492 (1992); Kantor et al., supra note 51.

The population may consist of interlocking neighborhoods, wherein individuals may be paired only with their neighbors. Interlocking neighborhoods were first considered in Glenn Ellison, Learning, Local Interaction, and Coordination, 65 ECONOMETRICA 1047 (1993). Ellison shows that this can significantly increase the rate of convergence, but has no real effect on the efficiency properties of the converged states. The legal literature on norms has focused on the case of interlocking neighborhoods. See, e.g., Picker, Single Games, supra note 4.

There are several possible interpretations of the matching process: (1) the agents are randomly paired once each period (perhaps with an odd man out), so that each agent expects to be matched with a given strategy according to its prevalence in the relevant population; (2) each agent actually plays each other in a "round robin" tournament all within the time frame of a "period," so that each player is actually matched with a given strategy according to its prevalence in the relevant population; (3) the agents are randomly matched as in (1), but are paired a large number of times within the period, so that the proportion of times that a player actually matches with a given strategy is nearly certain to be the population proportion of that strategy. The system is relatively robust to choice of interpretation, although the first interpretation precludes those yields different efficiency results if re-optimization proceeds purely by imitation and in gaming how well a given strategy performed in the previous period, agents do not account for how strategies were matched. See Arthur J. Robin & Fernado Vega-Redondo, Efficient Equilibrium Selection in Evolutionary Games with
particular, we imagine that each individual adopts a strategy for use against whomever she happens to meet in the current period. Occasionally she looks up from the process of playing this ongoing strategy, surveys the strategies employed by her fellow players in the last period, and begins playing the best strategy given this population. The re-optimization is naïve in the sense that the individual assumes that the current population composition of strategies is fixed for the foreseeable future, even though she herself is considering a change.

Mutation is modeled not as a genetic hiccup, but as experimentation, mistake (of calculation or implementation), the replacement of old players with uninformed entrants, or simply the noise of the great residual unmodeled. Importantly, mutation is an ongoing phenomenon, but nonetheless one that is rare relative to the process of occasional re-optimization described in the previous paragraph. Further, mutation is an individual phenomenon: the chance that one individual mutates does not depend on whether others also mutate.

It goes without saying that the evolutionary process thus described is only a mathematical parable. It is, however, a parable that has generated optimism about the possibility of widespread cooperation in human societies. We argue that this optimism is not fully justified, even when one accepts the premise that the mathematical representations of learning and evolution are helpful in understanding social interaction.

2. The Survival of the Inefficient

To see how this process would unfold over time, consider a simple evolutionary model based on the Stag Hunt depicted in Table 5. Imagine a population of ten hunters who go out separately every day to a favorite wilderness to hunt for food. On each hunting trip, each individual happens upon (exactly) one other hunter under circumstances described by the Stag Hunt: each faces the choice whether to join forces with her random partner to hunt the stag grazing just beyond the brush or else opt out for the hare running under foot. (We consider the possibility of choosing a partner below.) At any given time, each hunter follows a particular practice whomever she happens upon. Back at the hunting lodge, detailed records of each hunter’s kill are kept that communicate to each the manner in which her 9 fellows are currently hunting and, given this configuration of practices among her potential partners, whether her own current hunting strategy is still a good idea. For simplicity, let us
imagine that each individual takes at least one look at the hunting records per year, giving each at least one annual opportunity to re-optimize. (We can imagine that no two hunters re-optimize on the same day in order to make somewhat more reasonable the assumption that re-optimization proceeds under the naïve belief that opponent practices are fixed.) We will also suppose that on New Year's Eve each individual may "mutate" to whatever strategy she is not currently using with some probability set arbitrarily at 10% and drawn independently across hunters. 

Consider first the periodic re-optimization, ignoring mutation for the moment. Imagine that at the start of the year everyone hunts stag. At the first opportunity to re-optimize, a given hunter considers the fact that all 9 of her fellows are also hunting stag. She expects, then, that the next time she comes upon someone else on the hunt, this individual will be a stag hunter. Preferring half a stag to one hare, she resolves to continue her practice of cooperatively hunting stag. All

*. This part attempts to capture the key features of evolutionary game models in the simplest possible terms. This note examines the dynamics on a more technical level. The models in this literature are specifically designed to produce finite Markov processes, in which the state of the system in period \( n \) is a (probabilistic) function only of the state in period \( n-1 \), not of all history. Thus, given the current population composition of strategy choices, one can calculate the likelihood of every possible strategy configuration in the next period's population. The calculation is based on: (1) the probability that a given individual will "look up" from her current strategy and consider changing her behavior; (2) the imitation or optimization rule governing how individuals make this change; and (3) the probability (mutually exclusive from the chance in (1)) that each individual will mutate to each possible alternative strategy. The dynamic system is assumed to have a particular property called "ergodicity," meaning that, given any two possible strategy compositions for the population, there is always some chance, however miniscule, that the system could go from one composition to the other. This follows in turn from the auxiliary assumptions that mutation to any strategy is always possible for each agent and that mutation is independent across individuals. Ergodicity in the context of finite Markov processes also implies that in the long run, the possible states of the system can be summarized by a unique probability distribution over all possible population compositions called the "stationary distribution." NANCY L. STOREY ET AL., RECURRENCE METHODS IN ECONOMIC DYNAMICS 330 (1989). The probability assigned by the stationary distribution to any given population composition of strategies is two things at once: it approximates the expected chance that the population will be of that composition in the long run; it also approximates the actual expected fraction of the time that the population is indeed of that composition over the long run. The key technical results in this literature come from observing the asymptotic behavior of this stationary distribution as the probability of mutation is brought to zero. Both this limiting process and our simple calendar-year model capture the notion that mutation is far less frequent than re-optimization.
individuals re-optimize in the same way—without changing their strategies—and the population still consists of ten stag hunters. The same thing happens if we start with a population in which all individuals are hunting hare. When an individual re-optimizes, she realizes that she will be matched with a hare hunter, and preferring one hare to no food at all, she continues to opt for hare. With respect to the intra-year dynamics of re-optimization, then, these two social states—in which all hunt stag and in which all hunt hare—are "steady states": if society has arrived, by whatever means, at one of these social states, then re-optimization (by itself) will not carry it away. No social state other than these two has this property.\(^7\) We can therefore consider stag hunting and hare hunting as the two competing norms in this setting.

Now consider how the process behaves across the years, introducing mutation. The two steady states in which all hunt stag or all hunt hare are stable with respect to small numbers of mutations. Imagine that the population consists of all hare hunters and at year's end two individuals start hunting stag. Each other hunter will recognize that by hunting stag, she will achieve the cooperative solution if paired with one of the two mutated hunters. Under the circumstances, this would occur with probability 2/9 and would provide a daily payoff of 4. The expected daily payoff to hunting stag, then, is 8/9. However, because there are 7 other hunters hunting hare, her expected daily payoff from hunting hare is 2/9. Thus, as a stag hunter, the risk of encountering one of the more prevalent hare hunters and ending up with nothing is too great to justify a change of practice. Indeed, when few individuals who have mutated to stag hunting have a chance to re-optimize, they themselves will return to hunting hare. The population, having been altered slightly from the initial social state in which all hunt hare, will return to that state within the year. A similar story can be told for stag hunting.

Even though both homogeneous social states are immune to small

\(^7\) The stag hunt in Table 5, as a traditional game form, has mixed equilibrium in which 3/4 of the population hunts stag and the other 1/4 hunts hare. However, neither of these fractions can be expressed as a whole number part of 9. Thus, there is no social state in our population of 10 corresponding to this mixed equilibrium. This phenomenon is not coincidental—indeed, the opposite is coincidence. Almost all mixed equilibria in the class of all possible games will have irrational number break-even point, since almost all numbers are irrational. This means that almost every mixed equilibrium will be inexpressible as whole number parts of the population, no matter how large that population is. Nonetheless, the coincidental presence of a mixed social state would not materially alter the analysis here.
numbers of mutations, there is for each state some number of simultaneous mutations sufficient to "tip" the norm—that is, to cause the re-optimizing individuals to switch strategies during the course of the ensuing year until all have switched to the other norm. A mutation of 7 will tip the norm from hunting hare to hunting stag (assuming it will next be the turn of a nonmutated hare hunter to re-optimize; if one of the mutants is next in line, 8 mutations are required). Consider a hare hunter's optimization problem after 7 hare hunters have mutated to hunting stag. She now expects to encounter a stag hunter with probability 7/9. Stag hunting will lead to an expected daily payoff of 4(7/9) = 28/9, which is greater than 3, the expected payoff for hare hunting. After this individual switches to stag hunting, there will be 8 stag hunters and the next re-optimizing individual will follow suit by the same reasoning. By year's end, the entire population will be hunting stag. By contrast, 6 simultaneous mutations will not tip hare to stag. After 6 mutations, the expected payoff to stag hunting is 4(6/9) = 24/9 < 3.

While it takes at least 7 simultaneous mutations to tip the social state from hare hunting to stag hunting, going in the other direction is markedly easier. Imagine that 3 individuals mutate to hare hunting in a population that was formerly composed entirely of stag hunters. Assuming that a nonmutated stag hunter is the next to re-optimize, she can expect to come upon a fellow stag hunter with probability 6/9. The expected payoff to stag hunting is then 24/9, slightly worse than the guaranteed 3 of hare hunting. Hence, this stag hunter would switch to hunting hare. This decreases the chance of encountering stag hunters for everyone else, and by year's end the system will have tipped to a population consisting entirely of hare hunters.

Why is stag hunting so much more "dippable" than hare hunting? Compare stag hunting's efficiency advantage to its mismatch risk. Stag hunting does somewhat better against itself (payoff of 4) than hare hunting does against hare hunting (payoff of 3). The one-unit difference is stag hunting's efficiency advantage. But stag hunting does much worse against hare hunting (payoff of 0) than vice versa (payoff of 3). This 3-unit difference is stag hunting's mismatch risk. Because stag hunting's mismatch risk exceeds its efficiency advantage, hare hunters will not switch to stag hunting unless matching with another stag hunter is more likely than not—in our case, somewhat more than twice as likely.
3. The Dramatic Consequences of Small Differences in Tipping Thresholds

Given that hare hunting's tipping threshold is a bit more than twice that of stag hunting, one might imagine that the system would be twice as likely to bounce from stag hunting to hare hunting as vice versa. And, therefore, the system would spend somewhat more than two-thirds of the time at hare hunting and the rest at stag hunting. Not so. The time spent in each state will be dramatically more lopsided than the tipping thresholds.

The way we envision mutation (independent Bernoulli trials with $p = 0.1$) implies that the number of simultaneous mutations in a period has a binomial distribution. Consequently, the relative time the system spends in each equilibrium state varies exponentially, not multiplicatively, with the relative tipping thresholds. The number of simultaneous mutations required to tip from hare to stag (7) is $2 \frac{1}{3}$ times as great as the number required to tip in the other direction (3). However, the ratio of the probability of observing 3 or more simultaneous mutations to that of 7 or more simultaneous mutations is on the order of $0.1^{35}$, or 10,000, to 1.\(^{36}\)

These large differences in probability translate directly to differences of similar magnitude in the amount of time that the system actually spends at each steady state. Stag hunting renaissances are brief: if we start with a population of all stag hunters, there is a 90% chance that the system will have tipped to hare hunting within 3 decades. On the other hand, hare hunting periods last much longer: starting with a population of all hare hunters, there is a 90% chance that this medieval era will last at least 11,000 years and a 50% chance that it will still be in effect after 75,000 years.

Correspondingly, it takes much longer, on average, for the system to tip from hare to stag than the reverse, and we will accordingly observe stag hunting only during rare and fleeting periods. Interestingly, the smaller the probability of mutation in any period, the more pronounced the likely time difference.

C. Chaos in the Right Savannah

The possibility that evolutionary processes will select for the inefficient norm is driven in part by the risk to the efficient actor of encountering others not similarly disposed. But what if this risk could

\[^{36}\] Strictly speaking, the ratio is approximately 7095 to 1.
be managed? In particular, how does evolution proceed in settings where individuals have some choice over the individuals with whom they interact?

Ongoing research in evolutionary games has focused on some special forms of "endogenous interaction," under which evolution will select for the efficient strategy even when that strategy's mismatch risk outweighs its efficiency advantage.17 If we analogize changing networks to changing "locations," several ingredients are required for this result: (1) there must always be additional locations to which to move (by choice or mutation); (2) moving costs must be insignificant; (3) there must be no significant density effects (positive or negative—in particular, locations must not be capacity constrained18) and, on the other side, there must be no significant benefit to concentration; and (4) individuals must be fairly certain of the strategies used at all other locations.19

If these assumptions hold, contrary to the case described above, the inefficient norm will be more easily tippable than the efficient norm. To illustrate, suppose that the inefficient norm is being played in all populated locations. Formerly, tipping to the efficient norm (stag hunting) required 7 simultaneous mutations. Now tipping only requires that any two individuals "mutate" by simultaneously migrating to the same empty location and beginning to hunt stag. These two happy pioneers will then be doing better than the rest of

17 See, e.g., Jörg Oechsner, Decentralization and the Coordination Problem, 32 J. ECON. BEHAV. & Org. 119 (1997); Jeffrey Ely, Local Conventions (1998) (unpublished manuscript, on file with authors); George J. Mavitz et al., Endogenous Interactions (1997) (unpublished manuscript, on file with authors). For a dynamic pre-play communication model that operates by a similar dynamic, see YongGwan Kim & Joel Sobel, An Evolutionary Approach to Pre-Play Communication, 63 ECONOMETRICA 1181 (1995). Part of the emphasis on the restoration of efficiency here is no doubt due to the fact that these papers are meant to generate interest by being posed in opposition to the earlier work that derived the inefficiency result described above.

18 See Tore Dieckmann, The Evolution of Conventions with Many Players, 38 J. ECON. BEHAV. & Org. 93 (1999); A.W. Ansar, When Does Immigration Facilitate Efficiency? (1999) (unpublished manuscript, on file with authors). The relevance of these authors' results is limited by the assumption that when individuals are indifferent between locations, they move randomly according to a probability that is not taken to zero with the rate of mutation. This small assumption, which is unrealistic given any positive cost of moving, is a driving factor in some of the results.

19 Dieckmann, supra note 40, downplays the necessity of this assumption, but considers somewhat artificial forms of limited observability. V. Bhaskar & Fernando Vega-Redondo, Migration and the Evolution of Conventions (1998) (unpublished manuscripts, on file with authors), emphasizes that individuals must also be able to "optimize" location and action simultaneously. This seems to be a less-important factor in modification.
the population. Over the course of the ensuing year, whenever one of
the $8$ other individuals re-optimizes, she will pick up and move to the
new efficient frontier, changing her action in transit. By year's end,
the entire population will be grouped together in the new location
and the efficient norm will be universal. On the other hand, stag
hunting is not tipable by the same sort of spontaneous migration. If
two individuals leave a population of stag hunters and begin hunting
hare, no one will join them. Instead, the stag hunting must be tipped
in the old-fashioned way: $3$ individuals at the same location must
simultaneously switch from stag to hare. Since hare hunting is now
the more easily tipped, we can expect the system to spend a greater
proportion of time at the stag hunting norm.

Unfortunately, the list of reasons why this happy result is of
limited relevance is as long as the list of assumptions on which it is
based. First, there has to be somewhere new to go where the two
mutants can start fresh. If our two happy pioneers arrive at a location
that is already full of individuals hunting hare, they alone will change
strategies over the course of the ensuing year. Second, if the costs of
changing location (or social/business networks) exceed the efficiency
advantage of the efficient norm, then we would not see individuals
rushing to join the two relocated mutants on the frontier. Third, it is
clear that if individuals are unsure of whether things really are better
at the new location, they may be more likely to decide that moving is
not worthwhile. Uncertainty compounds the effect of moving costs
and can prevent the emergence of the efficient norm even when such
costs are relatively small. Last, individuals may decline to move if the
relative concentration of their present location affords network
benefits. Conversely, if locations become congested then the gains
from moving to the new location to join the two initial mutants will
crash as more and more individuals make the choice to move. At
some point, moving will no longer be worthwhile. Consequently, the
process will now stop short of universal efficiency (at the new
location) and play across the full population will be a mixture of
efficient and inefficient norms.

Perhaps the best assessment of the recent work on "endogenous
interaction" is that it tempers, to some degree, the pessimism of
earlier results that pointed toward inefficient outcomes when

---

42 Dieckmann considers the extreme case of congestion wherein locations are
capacity constrained. He finds that the system tends to either pure inefficiency or a
mixture of efficiency and inefficiency, but not to pure efficiency. Dieckmann, supra
note 40, at 170.
efficiency advantage is small relative to mismatch risk. Because of the very special structure necessary to restore efficiency in these endogenous interaction models, however, the pessimism is not so tempered as to become optimism. The models that restore efficiency rely on perfect information, the absence of friction, and location availability that runs counter to common experience and practical application.

III. MISMATCH RISK IN SIMPLE LEGAL GAMES

Mismatch risk, as we have demonstrated, reduces the likelihood that efficient norms will evolve. It may also help to explain why and where legal institutions are necessary. Our argument is straightforward: many of the standard problems to which law is addressed involve circumstances in which playing the efficient strategy requires an investment, whereas playing the inefficient strategy involves no—or a smaller—investment. Mismatch not only defeats the purpose of the efficient player’s investment, but may enable the inefficient player to appropriate some of that investment. Consequently, the efficient strategy involves greater mismatch risk than does the inefficient strategy. Depending on the relative size of the mismatch risk and efficiency advantage, this may make the efficient strategy less fit than the inefficient strategy.

To illustrate the point, we will interpret simple property, contract, and tort problems as analogous to the Stag Hunt game, in the sense that they produce two Nash equilibria and asymmetrical off-diagonal payoffs. This is not to say that these are the only, or even the best, ways to model these interactions in a 2 x 2 framework. One could tell the stories somewhat differently and produce different payoff structures. However, we believe that the key feature of our payoff structures—differential investment and consequent mismatch risk—is ubiquitous in the class of problems that fundamental legal rules address.

A. Property

Consider a conflict between two norms in a primitive tribe that inhabits an island. The first norm recognizes the first person to clear a piece of land as an owner who may exclude others. The second norm holds that every member of the tribe has an unfettered right to forage anywhere on the island.

The two norms can be mapped onto two strategies in a 2 x 2
game, which we will call "Plant" and "Plunder." A player Plunders by gathering any wild or cultivated plants she can obtain anywhere on the island. If both players Plunder, each receives a payoff of 3 units. In order to Plant, a player must invest 4 units in clearing, fertilizing, sowing, and so on. When matched against another Planter, each player may then harvest 10 units’ worth of grain, for a net payoff of 6. When matched against a Plunderer, by contrast, the Planter can harvest only 9 units’ worth of grain, the other 5 being lost to the Plunderer. Therefore, when there is a mismatch, the Plunderer receives 5 units and the Planter receives a net payoff of 1.

The matrix representation is as follows:

<table>
<thead>
<tr>
<th></th>
<th>Plant</th>
<th>Plunder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Plunder</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Table 6: The Plant/Plunder Game

Both (Plant, Plant) and (Plunder, Plunder) are Nash equilibria. However, in case of mismatch, it is better to be a Plunderer. Plant’s efficiency advantage (3 units) is less than its mismatch risk (4 units). Consequently, the system will be vulnerable to splitting from Plant to Plunder.

B. Contract

The decision whether to keep a promise maps easily onto a prisoner’s dilemma game. Imagine two people, each of whom has a good worth 3 units to her, but 5 units to the other person. Imagine further, however, that simultaneous exchange is technically infeasible, so in order to realize the gains from trade the parties must promise one another to deliver the respective goods. Moreover, each must send her good before she can verify whether the other has done so.

---

If both parties keep their promises, each gains 2 units. If neither keeps the promise, neither gains. However, if one keeps and the other does not, the one who does not gains 5 units and the other loses 3. The game is a prisoner's dilemma in which "Don't Keep" is a dominant strategy, as shown below.

<table>
<thead>
<tr>
<th></th>
<th>Keep Promise</th>
<th>Don't Keep Promise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keep Promise</td>
<td>2</td>
<td>-3</td>
</tr>
<tr>
<td>Don't Keep Promise</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Keeping Promises as a Prisoner's Dilemma Game

As we have shown, however, in a repeated form of the game in which parties choose strategies to be played consistently, the choice of strategies becomes a coordination game in which cooperation—here, a norm of keeping promises—is possible. Indeed, if we turn our attention from one-shot actions to norms, the analogy to the Stag Hunt is obvious. A norm of not keeping promises is similar to hunting hare. It permits only simultaneous exchange and foregoes the benefits available from longer-term cooperation, but it also avoids the risks inherent in trusting another. A norm of keeping promises, by contrast, is like hunting stag. It makes possible the larger payoffs possible from trust and coordinated activity. However, in the case of a mismatch, the promise-keeper, having relied on another's promise, will generally be worse off than had she stuck to simultaneous exchange. Moreover, it is often the case that the inefficient player can appropriate some of the efficient player's reliance investment. This makes it especially likely that mismatch risk will outweigh promise keeping's efficiency advantage, and consequently, keeping promises will be more tippable than not keeping them.

C. Torts

The decision to take reasonable care is also closely analogous to the Stag Hunt game. Consider the following stylized account of the interaction between two parties engaged in risky activity. Either party
can choose to take care at a cost of 3 units or not to take care. If both parties take care, no accidents occur and each party receives a gross payoff of 4 units (or a net of 1 after deducting the cost of care). If either party fails to take care, an accident ensues and each party receives a gross payoff of 0. Note that if one party takes care and the other does not, the accident will still occur and the party who took care loses the 3 units expended on care. The matrix representation below depicts this model. The decision to take reasonable care is simply the stag hunt game with all payoffs reduced by 5 units.

<table>
<thead>
<tr>
<th></th>
<th>Take Care</th>
<th>Don’t Take Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take Care</td>
<td></td>
<td>-3</td>
</tr>
<tr>
<td>Don’t Take Care</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Taking Care as a Stag Hunt Game

We could make the example still more pessimistic by assuming that a single player’s investment in care reduces the probability of an accident. In that event, the inefficient player would enjoy some of the benefit of the efficient player’s investment, making it likelier that the game would be a prisoner’s dilemma. Once again, under plausible parameters, the efficient norm will be more tippable.

Each of these examples shares a simple feature: the efficient action entails an expenditure or opportunity cost while the inefficient action does not. Those expenditures are wasted in the event of a mismatch. This feature, if nothing else, is realistic. The central justification for enforceable property rights is to encourage investment. A key insight in the literature on contracts is that enforcement enables parties to make investments in reliance on the contract. Tort law imposes liability in order to create an incentive to make cost-justified investments in care.

46 See generally GUIDO CALABRESE, THE COST OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 68 (1970) (arguing that one goal of accidents law is to encourage
D. The Repeated Prisoner's Dilemma—A Generalization

1. The Role of Mismatch Risk and Efficiency Advantage

The pessimistic results in our games stem from the fact that mismatch risk exceeds efficiency advantage. By contrast, Picier and Axelrod choose game forms in which efficiency advantage exceeds mismatch risk. Recall from Table 4 that, when paired with always defect in Axelrod's tournament, tit-for-tat's mismatch risk was 8 units, while its efficiency advantage was 578 units.

That dramatic imbalance is a function of the particular payoffs and discount factor Axelrod chose and is not inherent in the repeated game structure. Consider, for example, the prisoner's dilemma payoffs laid out in *Game Theory and the Law*, by Baird, Certner, and Picker.5

<table>
<thead>
<tr>
<th></th>
<th>Cooperate</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>-2</td>
<td>.10</td>
</tr>
<tr>
<td>Defect</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Another Prisoner's Dilemma

Suppose that we repeat this game indefinitely. Tit-for-tat will be an equilibrium, a norm, so long as the discount factor (the probability of the game continuing for another round) is at least 33%. Nonetheless, even with a discount factor as large as 56%, tit-for-tat's mismatch risk will outweigh its efficiency advantage. For example, the collapsed repeated version of Table 9, with a discount factor of 50% (calculated as in Table 4), is shown in Table 10.

---

Table 10: Another Repeated Prisoner’s Dilemma with Discounting at $w = .5$

<table>
<thead>
<tr>
<th>Tit-for-Tat</th>
<th>Always Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tit-for-Tat</td>
<td>-16</td>
</tr>
<tr>
<td>Always Defect</td>
<td>-6</td>
</tr>
</tbody>
</table>

In this version of the repeated prisoner’s dilemma, tit-for-tat’s efficiency advantage of 6 units is outweighed by its mismatch risk of 10 units. Therefore, even though tit-for-tat is an equilibrium, evolution leads to the norm of always defecting, not the norm of tit-for-tat cooperation. Individuals would be less likely to adopt the cooperative norm when almost everyone else in their neighborhood was doing the same. However, the mismatch risk inherent in playing tit-for-tat is such that it would take relatively few mutations to a norm of noncooperation to convince all other players to follow suit.

2. The Role of the Discount Factor

This example establishes that tit-for-tat may not be evolutionarily fit, even when it is an efficient equilibrium of the repeated game. In other words, the distinction between efficiency and fitness carries over from Fick’s coordination games to the more complicated repeated games analyzed by both Axelrod and Bendor and Swistak. The optimistic results in those analyses turn on the assumption that the discount factor, $w$, is sufficiently large. A large discount factor will magnify tit-for-tat’s efficiency advantage relative to its mismatch risk, thus creating a bias in favor of efficient outcomes.

To see the relationship between discount factors and fitness, consider first that tit-for-tat’s mismatch risk comes in the form of a one-time hit. When tit-for-tat meets always defect, tit-for-tat does worse than always defect in the first period only. Thereafter, both strategies defect against each other and receive equal payoffs ad infinitum.

Tit-for-tat’s efficiency advantage, on the other hand, comes in the form of an infinite stream, much like a perpetual bond. When played

---

*Note, in particular, that the key theorems in Bendor & Swistak, supra note 4, all contain the proviso that discount factors are sufficiently large.*
against itself, tit-for-tat generates a constant stream of cooperative payoffs. When always defect is played against always defect, it obtains a constant stream of defection payoffs. The period by period difference between these two streams constitutes the infinite stream of tit-for-tat’s efficiency advantage.

Thus, no matter how large the one-time hit of tit-for-tat’s mismatch risk, and no matter how small the per-period difference that represents tit-for-tat’s efficiency advantage, there will always be a discount factor large enough so that efficiency advantage outweighs mismatch risk. Tit-for-tat’s one time mismatch risk may, for instance, be 900 times larger than its per-period efficiency advantage. But if the future is only barely discounted, then after only slightly more than 900 periods, efficiency advantage will “catch up” with mismatch risk.

It is worth clarifying that this analysis of discount factors is distinct from the analysis one conducts to test whether tit-for-tat is an equilibrium of the repeated game—and therein lies the distinction between fitness and efficient equilibria in a repeated game setting. Tit-for-tat is an equilibrium only if it does at least as well as always defect when one’s opponent plays tit-for-tat. Always defect does better in the first round because being a lone defector is better than being a joint cooperator. Tit-for-tat does better in all subsequent rounds because perpetual cooperation is better than perpetual defection. Players must care enough about the future to make tit-for-tat’s future advantage decisive. As we have seen, however, tit-for-tat is evolutionarily fit relative to always defect only if players care enough about the future that tit-for-tat’s mismatch risk (the difference between the lone defector’s, or “meanie,” payoff and the lone cooperator’s, or “sucker,” payoff) is outweighed by its efficiency advantage (the difference between the discounted value of joint cooperation in all rounds, including the first, and joint defection in all rounds).

There is no necessary relationship between the results of the equilibrium analysis and the comparison of efficiency advantage and mismatch risk—the outcome depends on the relative magnitudes of the payoffs in each cell of the game matrix and the discount factor. But we can generalize to the point of saying that tit-for-tat will be an (efficient) equilibrium and yet evolutionarily less fit than always defect whenever

\[
\frac{(m-c)(d-s)}{m-s} \geq W \geq \max \left\{ \frac{m-c}{m-d}, \frac{m-c}{c-s} \right\},
\]

where \( m \times c \times d \times s \) are the payoffs in the bottom left, top left, bottom right and top right quadrants, respectively, and \( W \) is the discount factor. The left-hand inequality is the condition for tit-for-tat’s
mismatch risk to exceed its efficiency advantage. The right-hand inequality guarantees that tie-for-tie is an equilibrium. It is easy to check that this inequality is satisfied by the example in Table 10 with which we began this discussion.

IV. LAW AND MISMATCH RISK

What are the implications of our "guarded pessimism" about the emergence of efficient norms? In the situations that seem to matter most for social welfare—those that involve investment—efficient norms are particularly fragile because of their mismatch risk and the consequent possibility of tipping to inefficient strategies. The situation is clearly not hopeless because we observe societies that protect property rights, enforce contracts, and so on. Social norms, however, are not typically the only enforcement mechanisms at work in these critically important situations. Instead, actors external to the "game," such as courts, administrators, trade associations, and so on, play an enforcement role. This Part notes how such enforcers correct for the evolutionary disadvantage that efficient strategies face.

Our chief purpose is to challenge the literature’s optimism that decentralized evolutionary processes lead to efficient outcomes. Having thus opened up a space for laws and institutions in the pursuit of efficiency, our modest objective in this final Part is to examine the character and contours of that opening. We do not take on the far more ambitious task of explaining how and under what circumstances legal institutions arise to fill the gap. The endogenous evolution of legal institutions is a fascinating problem, but it is beyond the scope of this paper. We treat law as exogenous and observe that it is effective at inducing efficient play in circumstances in which norms cannot.

A. The Threat of Punishment

Examining the structure of off-diagonal payoffs helps to focus attention on two different strategies for addressing norm violations: punishment and compensation. In the game matrix, a chief source of tipability is the difference between the (low) payoff in the upper right-hand cell and the (high) payoff in the lower left-hand cell. This mismatch risk can be addressed by reducing the payoff in the latter (punishment), by increasing the payoff in the former (reimbursement), or both. The third form of retributive occurs when we take part of the defector’s payoff and give it to the cooperator (compensation).

The norms literature focuses on strategy choices that impose punishment in the form of loss of future interaction. This is
consistent with experience from everyday life, in which many norm violations are addressed through punishment. When a neighbor routinely allows her dog to dig in my garden, I may punish through retaliatory action. I may stop taking in the neighbor’s newspaper when she is on vacation or cease other forms of cooperative action. Other neighbors, alerted to the norm violation by me or by their own observation, may punish through gossip or snubs.

Punishment in the form of withdrawal of future cooperation is effective at making cooperative strategies an equilibrium. A credible threat to punish will keep the other party in cooperative mode. Punishment does not, however, solve the problem of mismatch risk. The cooperator/retaliator never comes out ahead when she meets a strategy like always defect. Withholding future cooperation is costly to the retaliator. It reduces the always-defector’s future payoffs, but at the cost of reducing the retaliator’s as well. Thus, it does nothing to balance the account that has tilted in the always-defector’s favor as a result of her initial defection. For a numerical illustration, compare the payoffs in the one-shot game in Table 9 to those in the repeated form of the game in Table 10, and note that the mismatch risk of 10 units is identical in both. To the extent, then, that the retaliator can say honestly, “this hurts me as much as it hurts you,” a norm of cooperation enforced by retaliation imposes significant mismatch risk on those who adopt it. When we introduce mutation—when it is possible that the cooperative equilibrium will be disturbed—the ability to retaliate does not guarantee that cooperation will survive."

**B. The Legal Solution**

Legal institutions solve the mismatch problem in a straightforward way, by forcing the defecting party to compensate the cooperating party. They do so from “outside” the game. We could, of course, imagine a norm (strategy) of compensation. Such a norm would require punishment of anyone who defects and then fails to compensate, and in that sense compensation could arise within the game. But if the underlying game form creates mismatch risk, this particular cooperative strategy will face the same hurdles as other cooperative strategies—it will fare poorly when paired with noncooperative (and noncompensatory) strategies.

This observation meshes with what we observe in the real world, where compensatory systems rarely arise in a completely decentralized

---

"We ignore here a separate and vexing problem: whether the threat of punishment will actually be carried out."
setting. Indeed, Ellickson comments on the fact that explicit compensation for losses was almost entirely absent from the system of informal enforcement of neighborhood norms. By adding the possibility of compensation, legal institutions reduce mismatch risk in a particularly cost-effective way. When we take part of the defector's payoff and allocate it to the cooperator, we double the reduction in mismatch risk compared to just punishing or to just reimbursing from a central fund. The legal system's punishments, therefore, are (dollar for dollar) twice as effective as decentralized punishments in reducing mismatch risk. This fits nicely with the problem that gives rise to the necessity of an institutional solution—the inefficient player can often appropriate part of the efficient player's investment, effectively doubling the mismatch risk created by that investment. A compensatory regime precisely reverses mismatch risk, making the efficient strategy robust to occasional mutation. In a world of costly collection, it also economizes on enforcement cost.

The surprise, then, is not that law solves mismatch risk. Once we allow for the possibility of an external actor to alter payoffs, there are unlimited ways in which the parties can be encouraged to play efficient strategies. Rather, the interesting point is that law solves mismatch risk with such economy. Consider other ways in which we could induce efficient play. In principle, either increasing the efficient strategy's efficiency advantage or reducing in mismatch risk are equivalent solutions. We could try to improve efficiency advantage by rewarding both players for playing the efficient outcome or punishing both players for playing the inefficient outcome. The former is simply not consistent. One must be able to collect in order to reward, but collection in an efficient regime would lower efficient payoffs and be self-defeating. Punishing when both are inefficient is consistent, but not as cost effective as a regime of compensation for mismatch. Punishing joint defection extracts resources from both players which can be returned to the system only if and when efficiency arises. In contrast, when extraction from the defector is combined with reimbursement to the cooperator to create

---

59 To illustrate the problem, note that an attractive means of reducing pollution in the domestic context is through pollution taxes that, in effect, require defectors (polluters) to compensate the public treasury, thus indirectly reducing the tax burden on cooperators. In the international context, however, the lack of an enforceable state makes this strategy impractical. Instead, developed co-entities have had to, in effect, bribe less developed countries to cooperate (that is, participate in pollution-reducing treaties). See Jonathan Raets Wiener, Global Environmental Regulation: Instrument Choice in Legal Context, 108 Yale L.J. 677, 785-87 (1999).

60 See ELICKSON, ORDER, supra note 5, at 60.
compensation, the resources collected from one player are immediately returned to the system, thereby reducing the overall cost of the enforcement strategy.

To illustrate these principles, we return to the Plant/Plunder Game from Part III.A. As we discussed in that Part, the feature of the game that makes Plant a less fit strategy is the 4-unit investment required to Plant, which leads to an off-diagonal payoff of 1 for a player who Plants but is mismatched with a Plunderer. The Plunderer mismatched with a Planter can appropriate some of that investment, receiving a payoff of 5. This mismatch risk—the risk that the 4-unit investment will be lost to the Planter and gained by the Plunderer—means that the burden of a mismatch (which may happen unexpectedly when mutations are possible) falls most heavily on the Planter. Selecting Plunder acts as an insurance policy that avoids the lowest possible payoff at the cost of abandoning the possibility of the highest possible payoff. Evolution consequently favors Plunder under many initial parameter values.

We can easily switch the direction of this effect, however, by switching the incidence of the mismatch risk. Imagine, for example, that whenever there is a mismatch, the Plunderer must compensate the Planter for the appropriated 4-unit investment. In that event, the game looks the same as in Table 6, but with the off-diagonal payoffs inverted, as shown in Table 11.

<table>
<thead>
<tr>
<th></th>
<th>Plant</th>
<th>Plunder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Plunder</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 11: The Plant/Plunder Game with Compensation for Lost Investment

Now the risk of a mismatch falls most heavily on the Plunderer, and evolution "dips" in the opposite direction. Indeed, it does so quite rapidly, as Plant is a dominant strategy.\[2\]

The same result holds when we look at a prisoner's dilemma game rather than a coordination game. Return to the Tree Maintenance

[2] Compensation cum punishment will not always make the efficient strategy a dominant strategy.
game of Table 1 and imagine that when only one party pays, the other must compensate for half of the cost (or $25). The off-diagonal payoffs now become (15, 15) and "Pay" becomes a dominant strategy.

We can generalize these results as follows: when the efficient strategy requires investment that is appropriate in the case of a mismatch, requiring the player who chose the inefficient strategy to compensate the other for the lost investment makes the efficient strategy the most evolutionarily fit. This specific form of payoff alteration is sufficient, but not a necessary, condition for efficiency. It is, however, a relatively parsimonious alteration that takes advantage of the "double effect" phenomenon described above and does not remove resources from the system for long periods.

CONCLUSION

Norms scholars have discussed a variety of means by which law can interact with social norms. Law might reinforce norms to the extent judges or legislators deliberately make legal rules consistent with extant norms. Law may provide a focal point that creates an expectation that other actors will play the focal strategy, thereby helping to establish a Nash equilibrium. Legal or other institutions might "seed" a norm by introducing cooperators into an evolutionary environment in hopes of moving the system from a less- to a more-efficient equilibrium. Law might also help a system from a bad norm to a good norm through a brief period of punishing players of the bad norm.

We propose a more substantial role for law and other institutions. Efficient norms can be surprisingly fragile in response to random shocks. Institutions provide a sense of security to cooperators that the cost of cooperation will not be wasted (nor compounded by the cost of retaliation). Institutions do not serve merely to get the system on the right path and then sit back and let it converge on its (efficient) course. Rather, they provide ongoing mid-course corrections when the system slips off the equilibrium path.

54 See generally Lawrence Lessig, The Regulation of Social Meaning, 62 U. CHI. L. REV. 943 (1995) (discussing solving coordination problems through mutually recognized focal points that allow expectations to converge and suggesting law as a means of creating a focal point and achieving coordination).
THE SHAREHOLDER WEALTH MAXIMIZATION NORM AND INDUSTRIAL ORGANIZATION

MARK J. ROE

Industrial organization affects the relative effectiveness of the shareholder wealth maximization norm in maximizing total social wealth. In nations where product markets are not strongly competitive, a strong shareholder primacy norm fits less comfortably with national wealth maximization than elsewhere because, where competition is weak, shareholder primacy induces managers to cut production and raise prices more than they otherwise would. Where competition is fierce, managers do not have that option. There is a rough congruence between this inequality of fit and the varying strengths of shareholder primacy norms around the world. In continental Europe, for example, shareholder primacy norms have been weaker than in the United States. Because Europe’s fragmented national product markets were historically less competitive than those in the United States, their greater skepticism of the norm’s value came closer to fitting the structure of their product markets than did any similar skepticism here. As Europe’s markets integrate, making its product markets more competitive, pressure has arisen to strengthen shareholder norms and institutions.

INTRODUCTION ................................................................. 2064
I. SCOPE ............................................................................. 2064
II. THE UTILITARIAN BASIS FOR SHAREHOLDER WEALTH MAXIMIZATION .................................................................................... 2065
III. SHAREHOLDER WEALTH MAXIMIZATION AND MONOPOLY RENTS ............................................................. 2066
   A. Shareholder Primacy Could Diminish GNP if Industry Is Concentrated ..................................................................................... 2066
   B. The Ambiguity of Pro-employment Corporate Governance Policies ............................................................. 2068
   C. Why Not Both? .................................................................... 2070
IV. DELIBERATELY WEAKENING SHAREHOLDER WEALTH MAXIMIZATION ......................................................... 2071
   A. Corporate Law’s Standards ................................................................. 2072
   B. Rhetoric and Culture .................................................................. 2073
   C. Hostile Takeovers .................................................................. 2074
   D. Incentive Compensation ................................................................. 2075
V. CAN STRONG CAPITAL MARKETS EASILY OVERCOME WEAK PRODUCT MARKETS? ................................................................. 2076

† Professor, Columbia Law School. Thanks for comments go to Ian Ayres, Einer Elhauge, and Louis Kaplow.

(2063)
INTRODUCTION

I make a single, simple point in this Article: the relative value of shareholder wealth maximization for a nation is partly a function of that nation's industrial organization. When much of a nation's industry is monopolistically organized, maximizing shareholder wealth would maximize the monopoly's profits, induce firms to produce fewer goods than society could potentially produce, and motivate firms to raise price to consumers beyond that which is necessary to produce the goods.

I. SCOPE

A few words on scope: I do not specify here the source of shareholder primacy norms. These norms could come from culture, from institutions, or from rules. Nor do I focus on what does the work: the norm or the underlying institutions and rules. I instead focus on the fit between the norm (or its underlying institutions and rules) and industrial organization.

Disentangling norms from practices is not easy. It would be odd for a nation to have norms sharply differing from practices, institutions, and laws. Because norms are usually congruent with practices, institutions, and laws, knowing which element is critical is hard.

Another limit to the paper: I have nothing here to say on the absolute value of a shareholder primacy norm. It may be so critical for organizing large private firms that even where it fits badly with industrial organization, it is nevertheless worth pursuing. Or it may so diminish total social wealth (a minority view today) that it is not worth pursuing single-mindedly anywhere. Or it may be so brutal if pursued single-mindedly that no organization and no society can absorb it in its pure form. Perhaps the norm must be softened to survive. I instead make the point that its relative value depends on industrial

organization and that there has been a rough congruence around the world with this relative value.

II. THE UTILITARIAN BASIS FOR SHAREHOLDER WEALTH MAXIMIZATION

The prevailing academic and business view in the United States is that shareholder wealth maximization fits with a utilitarian, greatest-good-for-the-greatest-number philosophy. But would a nation with concentrated industry be as well served by strong shareholder wealth maximization institutions?

Shareholder wealth maximization is usually accepted as the appropriate goal in American business circles. The norm makes some uneasy, though: after all, why should shareholders, who usually are favored members of their society, prevail over, say, current employees, who usually are less favored?

The utilitarian justification is that this preference is the price paid for strong capital markets, and allocative efficiency and that these benefits are so powerful that they overwhelm the normative benefit of any distributional favoring of current employees over current shareholders. In the long run, the argument goes, employees and other stakeholders are overall better off with fluid and efficient capital markets, managers need a simple metric to follow, and both wealth and, in the end, fairness are maximized by shareholders being the corporation’s residual beneficiary, with the other citizens getting what they want via contract with the corporation. Current employees might be made worse off in some industries, but employees overall will have more opportunities, higher salaries, and better working conditions. Furthermore, a stakeholder measure of managerial accountability could leave managers so much discretion that managers could easily pursue their own agenda, one that might maximize neither shareholder, employee, consumer, nor national wealth, but only their own.

But that sketch is weaker in a nation with mostly concentrated industry and weak product market competition. Enhancing

---

1 See, for example, the famous essay by Milton Friedman, The Social Responsibility of Business Is To Increase Its Profit, N.Y. TIMES, Sept. 13, 1970, § 6 (Magazine), at 33. Although aggressive when it appeared, Friedman’s perspective is now mainstream in American business circles and was not unthinkable then (as it was in some other nations). Cf Michael E. Porter, The Microeconomic Foundations of Economic Development, in WORLD ECONOMIC FORUM, THE GLOBAL COMPETITIVENESS REPORT 1999, at 33, 42 (“In western Europe . . . the inability to place profitability at the central goal is . . . the greatest constraint to economic development.”).
shareholder wealth maximization in that kind of a national economy may, even if the baseline utilitarian argument is correct, reduce national well-being, as we next see.

III. SHAREHOLDER WEALTH MAXIMIZATION AND MONOPOLY RENTS

A. Shareholder Primacy Could Diminish GNP if Industry Is Concentrated

Consider the monopolist's discretion. In Graph 1, a stripped-down version of the basic supply-demand setting for a monopoly, the monopolist can restrict production, raise price, and maximize its monopoly profit by finding the price-quantity combination that makes the 'rectangle' (and, hence, its profits) as big as possible.

The monopolist could, if it wanted to, produce the amount that a competitive industry would produce. If it did so, it would not be maximizing its own wealth, because it would be leaving that 'rectangle' of profits on the table. The consumers' triangle would be maximized. The monopolist would also, incidentally, usually employ more people as production would rise compared to the constrained production when monopoly profits are maximized. It would maximize the nation's wealth, not its own.

A strong shareholder wealth maximization norm will induce the monopolist's managers to lower production, to raise prices, and to lower employment. Weaker shareholder primacy norms could (plausibly but not necessarily) increase national wealth in a nation with highly concentrated industry.

Weaker shareholder primacy institutions could induce the firm to move down the demand curve, producing more and lowering price. This point is consistent with a standard view on agency costs: If unconstrained managers usually prefer to build larger firms, if they usually prefer to build new factories, and if they usually
prefer sales to profits—all of which are typically the core managerial agency costs to shareholders—then they are more likely in a concentrated industry with weak shareholder wealth maximization to travel down the demand curve by first producing more and then concluding that they cannot aggressively raise price. The weakly controlled managers could produce more national wealth than the tightly controlled ones.\footnote{Cf. Thorstein Veblen, The Engineers and the Price System 70-71 (1936) (asserting that a "free hand" for managers would increase production).}

To be sure, this increase in social wealth, while a plausible effect of weakened shareholder primacy in a monopoly, is not its only possible effect. Most obviously, loosening constraints on managers may just mean that the managers take more for themselves rather than increase production. One could, however, imagine a society with weak shareholder primacy norms but strong antitheft institutions: Managers are precluded from taking for themselves but are not instructed as to what they can do with the potential pot of monopoly profits. When incentives and markets are weak, they have discretion. Even here there is ambiguity. In the United States, "managers taking more for themselves" has led to higher salaries, more perks, and bigger but less profitable empires, all of which have generally negative social effects when product markets are competitive. But when product markets are uncompetitive, the firm that expands to produce more could be socially positive, assuming the empire is not an unproductive one.

Conversely, managers on a short leash might stay at the same point on the demand curve, but economize more on resources if they must maximize shareholder wealth. Economizing inputs tends to offset the maximizers' reducing output. In an economy with widespread monopoly, some firms encouraged to maximize shareholder wealth would primarily economize, while others would slash production and reduce allocative efficiency. One cannot predict which effect would dominate.

More subtly, ex ante incentives would diminish if, after a monopoly was acquired, institutions, rules, and norms weakened the shareholders' profits. Which effects—the negatives of incentive effects and managerial gains of the rents or the positive of a better ex post allocative efficiency—would dominate is a priori uncertain. Hence, nations with concentrated industries and many monopolies and oligopolies have reason to be less enthusiastic about, and
conceivably even to denigrate, shareholder primacy (depending on which effects—incentives and reducing waste or allocative efficiency—they predict would dominate). Historically, the ambiguous balance was strongly in play in continental Europe where monopolies were more widespread, but weakly in play in the United States where they were less widespread.

B. The Ambiguity of Pre-employment Corporate Governance Policies

Employees of monopoly firms can, and do, ally with capital to split the rents, to facilitate constricting production and raising price, and to seek barriers to competitive entry. That is, employees whose jobs are already set and secure often represent themselves, not the whole labor pool of all potential employees, some of those in that labor pool could be employed but are not, or are employed in less desirable jobs because the monopoly limits production and opportunity.

So, start here with the monopoly in Graph 1. Once the monopolist's employees gain a secure share of that rectangle, they become uninterested in policies that would move the firm down the demand curve to a lower price with more production because that would reduce the size of the rectangle (out of which the employees

\[\text{The proposition can be stated more formally. Imagine Nation A, a nation with competitive markets whose corporate law standard is that managers maximize national wealth. In a shareholder suit, managers concede that they were not managing shareholder profits. In Nation A's competitive market, they were therefore failing to maximize not only shareholder wealth but also national wealth. They should lose. Imagine Nation B, a nation without competitive markets, where the corporate law standard is that managers maximize national wealth. In a shareholder suit, managers must that they failed to maximize shareholder wealth, but were maximizing the firm's total sales, to the point where the revenues from the last sale equaled the costs of making that sale. Under a shareholder primary test, such managers would lose; but under a national wealth test they would win. Below, I suggest that it may be no accident that continental European nations have not had a strong shareholder primary standard. Infra Part IV.}\]
get some of their higher-than-typical wages and benefits) and, in turn, would thereby put downward pressure on the employees' wages and benefits. This much is understood.  

But consider the consequences of a supply shock to the industry. The cost of a critical input—oil, for example—rises dramatically. Usually the shape of the demand curve induces the monopolist to cut production further and lay off employees. That is, if a cost rises, the monopolist will often pass it, or most of it, along to the consumer by raising prices, cutting production, and laying off some workers.  

Consider the monopolist in a nation that denigrates shareholder wealth maximization and has rules and norms that discourage layoffs. Employees cannot easily be laid off. Their jobs cannot be radically reconfigured without their consent. Thus, the monopolist might not cut production and raise prices further, despite the shareholder-wealth-maximization bias for doing so, because it must pay the employees anyway if labor markets are rigid and if it cannot costlessly redeploy its workforce. In such circumstances, not only are the employees with jobs protected, but national wealth is increased (or at least not decreased) by slack agency controls on managers. A weak shareholder primacy norm facilitates greater production and, here, greater allocative efficiency. (To be sure, prior to the monopoly arising, employees with jobs at the potential monopoly firms would like the firm to have monopoly profits because then they can share in some of the rents. In this post-monopoly setting, they prefer that some of the monopoly profits be dissipated on keeping themselves employed. Ex post and ex ante attitudes to the monopoly results differ.)  

This is not just an abstract possibility. Detailed industry studies show that when German and American firms faced exchange rate pressures, German firms absorbed the cost, that is, shareholders

---


8 One could consider this scenario to be shareholder wealth-maximization inside constraints. But this scenario can play out in two ways: labor rules might constrain the owners from laying off the workers or corporate law and institutions could weaken shareholder-wealth maximization. The first is shareholder wealth maximization within constraints, the second is not.
shouldered the loss. Similar American firms passed it on, thereby risking that their higher prices would lead them to lay off employees and downszie, that is, employees bore more of the risk. Presumably the American firms did not absorb these costs because they were in a more competitive market and could not absorb the increased costs.

C. Why Not Both?

One could imagine a nation seeking shareholder wealth maximization within constraints. That is, the nation’s polity might constrain firms from laying off employees, but within those bounds, its rules and norms could allow or even encourage firms to maximize shareholder wealth in the belief that maximizing it will economize on resources and promote national wealth after basic fairness is assured. This combination is uncommon for at least three reasons. First, maximizing within constraints does not necessarily maximize social wealth when product competition is weak; some public policymakers may understand this. Second, bright-line fairness and job security constraints are often hard to write down in a law and enforce since the boundaries are hard to define: When are layoffs justified? When are they unjustified? Who judges when they are justified? Weak shareholder wealth norms and institutions plausibly could better operationalize the core goals of such a society that sought to favor employees. Third, political parties that constrain the firm (with pro-employee rules) often have complementary anti-shareholder ideologies; protecting jobs in place and denigrating shareholder value are an ideological (and public choice, interest group) package; political parties seek both, not one or the other, and, if a party achieves power, it implements both.

---

4 Michael Kneeter, Price Discrimination by U.S. and German Exporters, 79 AM. ECON. REV. 196, 205 (1989); cf. Michael A. Salinger, Tobin’s q., Extinction, and the Concentration-Profit Relationship, 15 RAND J. ECON. 159, 166 (1984) (concluding that either “there is little long-run monopoly power in the U.S. economy” or any such power is captured by an industry input such as unionized labor).

5 See supra note 8 and accompanying text.

6 High monopoly rents throughout the economy create two related cross-cutting effects: First, the potential for agency costs in the monopoly firms rises because the pot of monopoly money is up for grabs. Owners react by seeking to contain the profit loss, often keeping concentrated ownership because they cannot trust managers when monopoly profits make agency costs very high. Second, public policymakers who are looking out for the broad public have little reason to foster shareholder wealth maximization institutions.

8 One could similarly imagine less-competitive nations trying directly to force
D. Proviso: Who Pays for That Transferred Rectangle and That Lost Triangle?

Thus far I have focused on market structure in a national economy. In such a setting, shareholder wealth maximization norms and supporting institutions may decrease national wealth because the monopolist seeking to maximize profit would destroy more consumer value than the one with weak profit goals. For some small nations, however, the relevant consumer pool (whose wealth the monopolies are reducing) is not domestic but international.

So, now change our assumption of monopoly firms selling into a domestic market. Assume that the monopolistically organized industries export heavily to neighboring nations. Then some portion of the monopolist's rectangle comes from foreigners and some part of the diminished consumer's surplus triangle is lost to foreign consumers. If the policy choice is binary, the locally rational decision is to maximize monopoly profits at the expense of foreign buyers. Shareholder wealth maximization is not general wealth maximization (because foreign buyers are made poorer), but it is local wealth maximization.

IV. DELIBERATELY WEAKENING SHAREHOLDER WEALTH MAXIMIZATION

Corporate governance institutions tend to match the underlying organization of industry. Some of these institutions are legal, some economic, some cultural. A few examples follow of the differing strength of shareholder institutions.

monopoly firms to increase production. This they sometimes do, but usually via government ownership of the firm or its price regulation. Antitrust policy obviously plays a role here: when successful in anchoring competition (perhaps because the underlying conditions in the economy facilitate success), shareholders' primary aim play a larger role.

The point here is not that the antishareholder authorities are crisp in their thinking and policies. (And shareholder wealth maximization norms and institutions are also denigrated in nations with highly competitive markets, although denigrated less vociferously.) I do not think that public players have rigorously thought through the relationship, although a few surely have the intuition. The point is more that weak shareholder norms fit and perhaps are less likely to be challenged, in a weak competitive environment.

11 Sometimes the local monopoly can price discriminate by charging the monopoly price internationally and a lower price domestically.
A. Corporate Law’s Standards

Section 2.01 of America’s core academic aspiration for corporate law tells the firm’s managers: “a corporation . . . should have as its objective the conduct of business activities with a view to enhancing corporate profit and shareholder gain.”14 The analogous corporate law commands in continental Europe differ.15 French corporate law allows managers, it is said, to manage their firm in the social interest.16 German law refuses to tell managers that they are their shareholders’ agents. This tendency has been part of a long tradition of refusing to endorse shareholder primacy,17 and Germany has labor strongly represented on the large firm’s supervisory board. Indeed, German commentators (more authoritative there than are U.S. commentators here) state that German directors cannot act “only in the interests of the shareholders.”18 The fit of the differing national corporate law standards with the product market analysis thus far is obvious. I do not mean by this that corporate law’s instructions to managers determine what they do. The means by which managers can undo this instruction are many: the American business judgment rule vests them with enormous, nearly unreviewable discretion; monetary incentives, if lacking, could overwhelm any of law’s rhetorical instructions; and America’s corporate law itself has more

14 PRINCIPLES OF CORPORATE GOVERNANCE: ANALYSIS AND RECOMMENDATIONS § 2.01(2) (1994) [hereinafter PRINCIPLES OF CORPORATE GOVERNANCE] (emphasis added).

15 Philippe Biais, Les valeurs étroites du droit du ‘gouvernement de l’entreprise’ (The True Stakes of the Debate on ‘Corporate Governance’), ENVIR. DES SOCIÉTÉS, Jan-Mar. 1998, at 5, 15 (noting that, in France, as in most continental European nations, the social interest is the directors’ compass (quoting Alain Vandan, Professor of Law at the Université de Paris 2)).

16 French law, judges, and CEOs use the notion of ‘social interest’ differently. The judges [use it] to guarantee the ‘continuity of the firm, especially when [it] faces economic difficulties, [while] the CEOs refer to it to keep a free hand in managing the company. Shareholders see [it] as an ambiguous notion, mostly used against their own interests.


than a little undertow via its authority to account for ethics\textsuperscript{9} and, in reaction to hostile takeovers, to account for stakeholder constituencies. International differences in corporate law’s standards are the beginning, not the end.

B. Rhetoric and Culture

The rhetorical pressure goes beyond law. Norms in American business circles, starting with business school education, emphasize the value, appropriateness, and indeed the justice of maximizing shareholder wealth (which will trickle down, or raise the tide that will raise all boats, etc.). In France and Germany, shareholder wealth maximization is demeaned and seen as at odds with social values. And in Japan, senior managers rank shareholder profit maximization (more precisely: return on investment and stock price) much lower than do American managers.\textsuperscript{39}

Well-to-do families in some nations are said to prefer family ownership of enterprise. Firms are passed from generation to generation; corporate governance sometimes becomes the governing of family relationships. While this cultural preference may well be a fully independent dimension to the organization of business and politics, note the fit between this cultural preference and high rents. The preference is “functional” (for shareholders) in that such family structures can keep more of the rents inside that family than can another structure. It can be seen as a natural reaction when shareholder wealth institutions are weak, making a separation of ownership from control more precarious for shareholders than when they are strong. (Or, reversing causality but keeping the fit: weak shareholder primacy institutions may fit when ownership is close anyway, as the dominant stockholders can look out for themselves.) An American culture of a founder’s quick sale of her successful firm fits better with an economy with few rents (and with stronger

\textsuperscript{9} See PRINCIPLES OF CORPORATE GOVERNANCE, supra note 14, § 2.01(b)(2)-(3) & rep. note 5.

shareholders primary institutions and others).

C. Hostile Takeovers

In the 1980s, about thirty percent of America’s Fortune 500 companies received takeover bids.24 This is an extraordinary number, indicating that shareholder power via takeover bids had to be on the minds of all large firm managers. The 1980s were also, consistent with the thesis here, arguably one of the periods of strongest product market competition. Not only were American manufacturing markets workably competitive, but international competition was, for essentially the first time, pounding every manufacturer that could not perform. Hostile takeovers were, and despite the rise of the poison pill still are, an engine of shareholder wealth maximization.

In Europe until recently, hostile takeovers (and indeed any takeovers) were destabilized. The few hostile takeovers tried in Germany founded (until the Vodaphone takeover in 2000), often due to political pressure, as workers campaigned to block the takeovers and politicians sided with employees and against the capital owners.25 In one major attempt, in the steel industry, the nominally conservative German Chancellor said he was "deeply concerned" over it, asking the firms and players to "live up to their social responsibilities."26 They substantially cut back their planned restructuring.

The French Minister of Finance has been suspicious of high-priced takeovers because, he said when denouncing one such high-price offer in 1998, the "high [stock] price means the buyer would have to look immediately at: higher profits to pay for the acquisition, which


25 See Richard Cateline, Steel is Put in the Swoop, INDEPENDENT (London), Mar. 23, 1997, at 3 ("The combined forces of federal economics minister Guenter Rexrodt and Johannes Rau, premier of North Rhine Westphalia, Germany's biggest state, bounced Germany's virtually sole corporate bidder into 'negotiations' with Thyssen [for a smaller merger on terms favorable to incumbent employees.]"); Steel for a Battle, FIN. TIMES (London), Mar. 23, 1997, at 9 (reporting a mainstream German newspaper headline asking if an executive seeking a takeover wanted to set Germany on fire); Michael Woodhead, A Paralytic Victory for Germany, SUNDAY TIMES (London), Mar. 26, 1997, § 3, at 1, col. 1 ("The collapse of Krupp's bid for Thyssen is a victory for the social consensus.").

26 William R. Emmons & Frank A. Schnitz, Universal Banking, Control Rights, and Corporate Finance in Germany, FED. RES. BANK ST. LOUIS REV., July/Aug. 1998, at 19, 22 (quoting German Chancellor Helmut Kohl).
could be negative... for jobs.

Until the late 1980s, the state often decided takeover results and, even when it withdrew from overall control, it continued to seek to avoid takeovers that would yield "a social massacre" with "massive layoff[s]." The French ministers proposed a takeover law in March 2000 that would require an offering company to agree on some terms with the employees of the target. "A takeover could not succeed without taking into account employees' views," said the French Finance Minister, seeking to formalize what had been an informal policy.

Only recently, as European governments have been moving to the right economically and product markets have become more competitive, have hostile offers appeared; historically they occurred in continental Europe at a rate far lower than that prevailing in the United States.

D. Incentive Compensation

Stock-based incentive compensation could induce managers to maximize shareholder wealth. Studies of incentive compensation show that French, German, Italian, Swedish, and other continental European managers receive lower incentive compensation than American managers and that the incentive compensation they do receive is typically a much smaller fraction of their total pay.

Stock options have been less widely used in Europe. Partly this has been because stock options were disfavored in tax terms, and presumably much of this tax result came from an anti-shareholder perspective. Stock options have also been disfavored because options would further separate managers from employees, something that European culture sees as having an unethical edge.

---


V. CAN STRONG CAPITAL MARKETS EASILY OVERCOME WEAK PRODUCT MARKETS?

One might think that capital markets must induce shareholder wealth maximization even if product markets are weak. This is incorrect, however.

A simple explanation is that the two tend to move together, so disjunction (one strong, the other weak) is an odd result. But even if capital markets are strong and product markets weak, capital markets probably would not trump the weak product markets and induce strong shareholder results.

A. Raising New Capital

Capital markets’ constraints on managers are weaker when product competition is weaker. Consider how capital market competition constrains managers: Managers must go to capital markets for funds and, when they do, stock-buyers penalize poorly performing managers by demanding a higher rate of return and a lower stock price; creditors penalize those managers by demanding a higher interest rate; and at the limit, capital-providers refuse to give those managers any new capital. In the latter case, the firm withers. More effective firms with more effective managers eventually replace it.

But the capital market constrains the monopolist’s managers less strongly than it constrains a competitive firm’s managers. Often the monopolist’s managers can generate sufficient profits internally to pay for needed capital improvements. And as long as they leave some of the monopolist’s “rectangle” on the table for the original capital providers, the monopolist’s return on invested capital will still be higher than that of a competitive firm. Capital markets constrain the monopolist’s managers strongly not when the managers are dissipating any monopoly profits, but when the managers go further and dissipate so much that the return dips below the competitive return for capital. Until then, managers are constrained only weakly by capital.

considered entirely ethical”); see also Graham Bowley, Vonchlet Launches Stock Option Scheme, FT INTL TIMES (London), Sept. 15/16, 1997, at 19 (“Only a handful of Germany’s biggest companies have adopted share option schemes, which differ from those in the US and UK because of strict German regulations on employee share ownership.”).

The capital markets mechanism would, in textbook fashion, be this: When the firm goes fully public, investors will capitalize the firm’s expected cash flows. Those cash flows will be the competitive return, plus the additional monopoly profits before
SHAREHOLDER WEALTH MAXIMIZATION 2077

markets.
Thus, if the monopoly rate of return is, for example, 30% when worldwide capital markets demand a 10% rate of return (for this type of company, this class of risk, etc.), then capital market constraints might strike the firm dead if managers throw away enough that the expected return declines below 10%. But capital markets will not directly constrain the managers as they take, share, or squander 20% of the monopoly’s return. That 20% cushion is the monopolist’s rectangle, the potential excess profits that create the potential for slack.

B. Takeover Markets

Takeover markets constrain managers, but they do not do so fully. Even when they add to constraints on managers, they do not substitute fully for competitive markets.

Consider the typical premium in takeovers, which approximates 50% of the pre-takeover trading price of the stock. Postit that competitive markets ordinarily keep managers from straying more than 25% away from maximizing shareholder profitability. That would mean that when markets are un-competitive, there is an additional 25% of slack in shareholder wealth that managers are accorded, before takeover kicks in, to constrain them from further straying from shareholder profitability. More plausibly, we have a series of overlapping, but partial, constraints; removing one increases the slack in most, although not all, firms. For some firms, takeover markets constrain as strongly as does product market competition, making each substitute for the other. But for other firms one constrains more strongly, such that loosening the stronger constraint does cut managers more slack.

Moreover, these same societies that have weak product market

agency costs, minus the monopoly profits lost due to agency costs. The original owner reaps the expected gains from the monopoly and any expected losses from managerial agency costs. Unexpected reductions in agency costs will be captured by future shareholders. And if the potential agency costs are high for a monopoly firm, then the original owner will often decline to allow the firm to go public so as to avoid the expected loss.

So, if the competitive rate of return is 10%, but the firm earns $30, not $10, on its investment of $100, the original owner will be able to sell the firm for $300, if agency costs would be nil. If agency costs would diminish earnings to $25, then the original owner will be able to sell for $250. If the original owner values the gains from going public on diversification, liquidity, change of work plans, etc.) at less than $50, he or she will keep the firm private.
competition have reason to be reluctant to facilitate those takeovers, which would tighten up shareholder profitability, plausibly inducing higher prices and lower production. Many of these countries have discouraged takeovers.\footnote{See supra Part IV.G.}

C. Capitalization of Monopoly Profits

Capital markets will bid up the stock price until the return equals the risk-adjusted competitive return. Posit that the competitive, risk-adjusted rate of return is 10% annually. In competitive industries, $100 of investment will return $10 each year. A monopolist builds, with a $100 investment, a monopoly that yields $50 annually. When the monopolist sets the firm to buyers who expect the monopoly to be retained, the buyers will pay $90. The original monopolist captures the $200 "rectangle." If agency costs are expected to diminish the firm's profitability to $25 (i.e., leaving $15 of monopoly profits for capital-providers but dissipating the other $5), then outsiders will pay $250 for the firm, and the original monopolist can capture only $150 of the "rectangle."

True, managers in these now diffusely held firms have the usual reasons to want to increase the firm's profitability from $25, to the $30 level that is attainable. They might do so. But capital markets do not necessarily force them to: The managers can raise capital, if they need to, so capital markets' strongest weapon does not punish them. The typical 50% takeover premium needed meant that an outsider would probably not mount a takeover (even if the society allows one), because the slack only justifies a 20% premium.\footnote{A third theoretical capital market constraint is worth mentioning: When the monopolist leaves the firm, she could conceivably "enrich" her (nearly) entirely with $350 of debt, with an expected return of $30. Cf. Michael C. Jensen, Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, 76 ANN. ECON. REV. 253 (1986) (explaining that, according to the "control hypothesis" for debt creation, managers are constrained by the reduction of free cash flow resulting from the issuance of debt). The constraint on managers is that they must then scramble to meet the interest payment of $30. To do so, they must capture the monopoly profits.}

VI. Three American Implications

One: American shareholder primacy institutions—always more or less strong—strenthened further when competition intensified in recent decades.
Two: What if the American economy is changing? What if monopoly, say in new technologies, is becoming more widespread? There is a trade-off, as there is for patents and monopoly analysis generally, of ex ante incentives versus ex post allocative efficiency. If markets are competitive, there is good reason to weight ex ante incentives much more heavily than ex post allocative efficiency. And we weight ex ante incentives heavily in the United States—for example, patent protection for twenty years. But we do give ex post allocative efficiency something on the scale: the patent after all does not go on in perpetuity. Some of this ex ante incentive versus ex post allocative trade-off may, if monopoly once again becomes important, spill over into debates about shareholder norms. When even patent-holding monopolies face some, albeit weak, competition, the trade-off in favor of ex ante incentives is easier to make than when the ex post allocative costs are very high.

Three: although I have here focused on one relationship—high monopoly rents fit less well with shareholder primacy than does a competitive product market—there is much more to the analysis: high rents also affect politics and corporate structure through other channels.

First, they affect corporate structure by raising managerial agency costs. They raise managerial agency costs because there is more for managerial agents to lose for shareholders, and several constraints on managers—product market constraints, obviously, but also usually capital market constraints—are weaker in monopoly settings. Moreover, employees and other stakeholders increase their demands on the firm, managers could more readily accede because the pot to divvy up is bigger, and these demands should further raise managerial agency costs. This pressure on the public firm via heightened agency cost helps to explain why there were fewer public firms in continental Europe. And, because the United States has historically been more competitive, it might help to explain the (relatively) easy time the public firm had in developing in the United States. As Europe’s product markets have become more competitive, its demand for, or at least tolerance of, shareholder primacy institutions has also increased. This tolerance has made the public firm more plausible than it was previously.

Second, these rents can affect political structures. Conflicts in divvying up those monopoly profits can be widespread in a nation where monopoly is widespread. If widespread, these conflicts can spill over into politics, fuel political ideologies and political parties, and
privilege quick government-led conflict resolution so that production can go forward. When an economy’s product markets are more competitive, these conflicts diminish. Hence, the consequences of the conflict—political spillover, ideologies, and high demand for means that reduce conflict—also diminish.

I further develop these effects of high monopoly rents on the corporation and on disfavoring the diffusely owned public firm (by raising managerial agency costs) elsewhere.29

CONCLUSION

Maximizing shareholder wealth where competition is weak, therefore, could plausibly reduce production, raise prices, and lower national wealth, especially if managers when unconstrained value production, sales, and expansion over shareholder profits, as American agency-cost analysis usually concludes. (Or, because shareholder wealth maximization norms and institutions also induce economizing on resources and other positive incentives, where competition is weak, the norm does not raise production as much, does not lower price as much, and does not raise national wealth as much as it does where competition is strong.)

Hence, where industry is weakly competitive, shareholder wealth maximization norms and institutions are relatively less effective in raising social wealth than they are in more competitive economies.30 And it would be a mistake to conclude that strong capital markets can trump weak product markets: if there is a worldwide return demanded by fluid capital at, say, 10%, then the monopolist who can potentially earn 30% has 20% of slack to keep for itself, to squander, or to share with other players in the system. The strongest capital market constraints do not kick in until the monopolist’s managers squander, steal, or give away two-thirds of its profits.

Europe has tended both to have been less competitive historically and to have had weaker shareholder primacy norms and institutions than the United States has had. American shareholder primacy institutions—never especially weak in the first place—strengthened considerably as product market competition became more severe in recent decades. And in continental Europe, where national

30 At least in the short, and maybe the medium, run. See supra note 5.
economies tended to have industry much more concentrated than it was in the United States and where pro-shareholder institutions tended to be denigrated, their weak shareholder institutions fit with their product market structures.\textsuperscript{16} Whether or not Europe’s past of social democratic, antishareholder ideologies and institutions were wealth-maximizing (or the degree to which they were) depends on how the trade-off was made between gains from higher production and losses from more slack (ex post in weak controls and ex ante in weakened incentives), and on how uncompetitive its product market structure was.

\textsuperscript{16} If information technologies create natural monopolies in the United States in a wide patch of the economy, the tight American fit of strong product competition and strong shareholder institutions could change. Cf. Paul Krugman, \textit{Unsound Advice}, N.Y. Times, Oct. 22, 2000, § 4, at 15 ("The inevitability of monopolies in a knowledge economy... creates new puzzles for antitrust policy. The Microsoft case poses real dilemmas, and it is surely only the first of many.").