1993

Irreversibility and the Law: The Size of Firms and Other Organizations

Saul Levmore

Follow this and additional works at: http://chicagounbound.uchicago.edu/journal_articles

Part of the Law Commons

Recommended Citation

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
Irreversibility and the Law: The Size of Firms and Other Organizations

Saul Levmore*

I. INTRODUCTION .............................................. 333

II. IRREVERSIBILITY AND ITS LEGAL SOLUTION ..................... 336
   A. Managerialism, Over-Expansion, and Preventive Measures ..... 336
   B. Mitigating the Irreversibility Problem ......................... 339
       1. The Historical Example of the Ultra Vires Doctrine ...... 339
       2. Obstacles to Corporate Expansion ......................... 340
       3. Delaware's Anti-Takeover Statute ......................... 345

III. AGENCY COSTS AND LEGAL COSTS ............................. 346
   A. Law as a Source of Transaction Costs and Irreversibility .... 346
   B. Reinterpreting The Nature of the Firm: Relative Agency Costs 349
   C. Predicting Firm Size ........................................ 353

IV. THE IRREVERSIBILITY PROBLEM IN OTHER ORGANIZATIONS ........ 354
   A. Other Asymmetries ......................................... 354
   B. The Irreversibility Theme in Non-Profit Organizations ...... 356

I. INTRODUCTION

In his celebrated article, The Nature of the Firm, Ronald Coase explains the existence and scope of the firm as reflecting a balance between the relative advantages of internal and external arrangements. Internal to the firm, the entrepreneur who coordinates and integrates activities avoids the transaction costs associated with forming, adhering to, and enforcing external, formal contracts. Talented managers can shift resources from one internal activity to another with relatively little friction. On the other hand, external arrangements provide the entrepreneur with clear messages, primarily in the form of prices. The entrepreneur who acquires supplies from external sources has an easy time substituting among inputs according to the relative prices of these supplies. The implicit costs of factors generated internally

* The Class of 1962 Professor and Barron F. Black Research Professor, University of Virginia School of Law. I am most grateful to Hideki Kanda, Glen Robinson, Roberta Romano, and Robert Scott for comments on an earlier draft.

are far more difficult to discern. It seems useful, for example, to think of entrepreneurs as regularly borrowing from one part of a firm to invest profitably in another division. One external alternative is to borrow from a more formal creditor. This alternative offers explicit interest rates (which enable the decision-maker to assess the wisdom of the investment), but often comes with constraints and a variety of transaction costs. In some sense, then, the picture is one of the firm as a planned economy. As the size and scope of an operation increase, central planning is more difficult and less efficient than market transactions because planners must develop their own sources of information, while markets are animated by prices and other signals. This insight, in which organizational size is seen as a function of competing transaction-cost considerations, is rightly regarded as one of the most important intellectual developments in economics, organizational theory, and law.

This explanation of the size and scope of the firm is in the form of a deep, suggestive framework. It can hardly be claimed that all internal arrangements are more flexible and less informative than all external contracts. Thus, a labor-union contract may curtail an entrepreneur's freedom to substitute one method of production for another, even while it makes the cost of labor quite plain. One might insist that such a contract is external rather than internal, but it does not diminish Coase's insight to note that some internal contracts are sufficiently explicit and rigid, thus making internal growth no better than (and perhaps inferior to) external arrangements. Conversely, some external contracts, such as relational contracts calling for one party to supply as much of a given good as another requires at externally determined or profit-sharing prices, resemble internal arrangements in so many ways that it seems pointless to insist that the internal-external dichotomy explains all organizational decisions with a single theme. Contractual relations are too rich for such a simple sketch. Nevertheless, the picture of a firm as something with boundaries that represent a balance between transaction costs and valuable information has proved to be worth several thousand words.

My primary goal in this Article is to enhance our understanding of the tradeoff between internal and external arrangements with a new theme or variable. The foundation of this theme is that there is a kind of ratcheting, or irreversibility, in the evolution of a firm such that it is easier to expand than to contract. This asym-
Irreversibility and the Law

Irreversibility draws attention to two distinct legal phenomena. First, in anticipation of irreversibility, there may be precommitments—in the form of laws—against (seemingly optimal) internal growth. As we will see, law may offer a means of guarding against growth that proves too great once irreversibility is taken into account. Second, law may itself be the cause of irreversibility, because interest groups or considerations of efficiency may generate legal rules that make contraction more difficult than expansion. There is a tension between these two roles. One describes law as part of a strategy for creating firms of optimum size, while the other depicts law as an impediment to natural, efficient firms. Part II develops a more optimistic story in which law can play a remarkably positive role. Part III begins with the easier argument: Laws create transaction costs that in turn affect the nature and scope of firms. The argument is not, however, a mere affirmation of Coase’s theory, because I introduce and emphasize the irreversibility problem. Moreover, I argue that many transaction costs imposed by law do not affect the size of the firm, or the internal-external tradeoff, because it is relative (inter-firm) transaction, or agency, costs that matter in understanding the choice between internal growth and external arrangements. I suggest that this reinterpretation of The Nature of the Firm draws attention to competition among firms with different ownership structures and agency costs. Moreover, it strengthens the case for the importance of the irreversibility idea as a driving force in some areas of law. Finally, Part IV looks to other kinds of organizations, including nations and eleemosynary entities for different examples of irreversibility and its antidotes.

4. Economists are familiar with the problem or possibility of asymmetrically flexible, or “sticky,” wages and prices, and there is some literature on a kind of stickiness in organizational growth and formality. See Jeffrey D. Ford, The Occurrence of Structural Hysteresis in Declining Organizations, 5 Acad. MGMT. Rev. 589 (1980). Irreversibility is too strong a word for this feature of organizational evolution, because the basis for my argument is not that growth is really irreversible, but rather that it is more costly to contract after internal growth than it is to sever external relationships. Still, I use the term irreversibility, rather than asymmetry or ratchet effect or stickiness, because these other terms come with the baggage of distinct phenomena.

5. See infra text accompanying notes 11-14.
6. See infra text accompanying notes 11-43.
7. See infra text accompanying notes 44-46.
8. See infra text accompanying notes 47-55.
9. See infra text accompanying notes 56-58.
10. See infra text accompanying notes 59-67.
II. IRREVERSIBILITY AND ITS LEGAL SOLUTION

A. Managerialism, Over-Expansion, and Preventive Measures

A familiar theme in the theory of firms and bureaucracies is that managers, if left to their own devices, may prefer to build empires—to extend their control over increasingly large budgets and activities, rather than to maximize profit, quality of service, or some other good that their principals would prefer. In the case of government activities, there may be little else that is quantifiable for managers to maximize (although there is the modern possibility that managers will try to maximize and advertise the budget cuts they can conceive and carry out). But in the case of profit-oriented firms, it appears that it is either quite difficult to design appropriate incentive-compatible rewards for most managers, or that the market is slow to develop such compensation patterns. Managers are widely thought to defend their positions and compensation against outside acquirers who, in turn, are generally thought to be controlled by other managers who seek larger empires for themselves. This is hardly the place to examine either the evidence for these propositions, or the problems principals face in designing and enforcing compensation schemes that align their agents' interests with their own.1 But it is useful to note that there is a sense in which the nature of the firm itself causes the empire-building problem. To the extent the size of the firm is limited because managerial flexibility and command are eventually overwhelmed by the need for (external) explicit prices, good managers will in part be those who can direct larger empires without explicit external signals. Many managers will wish to appear to be good managers by building large enterprises around themselves. If good managers are those who command and control large organizations (within which their talents substitute for explicit prices), then mediocre managers will generate a kind of “lemons problem” by also constructing large organizations, or at least attempting to do so. Sole proprietors may have an incentive to balance internal flexibility and external signals in a way that optimizes the size of their firms, but the nature of firms run by agents may be to imitate the size and scope of firms run by the most valued entrepreneurs.

Competition and a harsh environment in which inefficient firms fail can surely force managers away from selfish empire-building and toward transaction-cost-reducing steps. Nevertheless, one straightforward and perhaps superficial reaction to the tradeoff put forth in The Nature of the Firm is to predict that, as a positive matter, firms will tend on average to be somewhat larger than one would think by simply balancing the transaction costs and information advantages associated with external arrangements. This is because competition is unlikely to be so nearly perfect as to eradicate completely the managerial preference for size. The point is not simply to recall the familiar possibility of managerialism, or empire-building, as opposed to profit maximization,18 but rather to see the very idea that optimal size


12. See, e.g., F.M. Scherer, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 34-37 (1980) (collecting various sources). There is also the possibility that managers pursue growth strate-
and scope depend on managerial talent suggests that imitation of this ideal firm (by more mediocre managers) may be an important determinant of firm size.

If we can accept the notion that many managers will push for increased size and scope—for their own sake, because compensation is mechanically or even partially linked to size, or because managers attempt to resemble good managers—and that this preference may create a kind of stickiness, or irreversibility, because if things do not work out well for the enlarged firm, managers will not be eager to contract the firm, then the scene is set for the crux of the argument. The next step is to ask what principals, or even efficiency-minded social engineers, would do to avoid the costs of empire-building and over-expansion. Knowing that their agents will try to take the firm on a costly and perhaps irreversible path, what can principals do in advance to prevent this possibility of inefficient internal overgrowth?

Before answering this question, it may be useful to note that there are complementary ways of expressing the likelihood of irreversibility. Perhaps the most intuitive of these is to say that as a psychological matter, managers will find it easier to hire workers and to open new plants than to close down operations, abandon communities, and discharge workers who have become part of the family and culture of the firm. An alternative version of the description leans more on self-images than on interpersonal sensitivity; managers may see themselves as failures both if they stand still and do not tackle new projects, and if they abandon a project. Either way, it seems likely that there are tendencies in favor of internal expansion, and that these are not matched by rewards for efficient contraction. Moreover, the asymmetry seems universal enough to say something about underlying human tendencies: Generals do not like to retreat; sports leagues often add, but rarely drop, franchises; religions convert more quickly than they excommunicate; hospitals build and expand, but rarely sell buildings; nations invade and explore for new property, but rarely offer independence or suggest partitions, unless they are under great pressure to do so; and universities invest in new schools and institutes and increase their student and faculty populations, but resist shrinking, except perhaps in times of serious financial distress. It seems unlikely that in all these cases the principals’ goals (or social welfare) are consistent with frequent expansion and extremely rare contraction.

One solution to this agency problem must be for the principals to make it difficult for their agents to expand. If principals know their agent has more of an incentive to expand than they wish the agent had, and if they know that this overexpansion is costly because it is at least partly irreversible (if only because the agent will resist contractions), then principals can try either to hold back from their agent the power to expand the organization and tasks they have delegated, or try to establish mechanisms that resist expansion or encourage contraction. Equivalently, as the principal-agent literature has shown, agents might themselves put such “bonding”

gies to achieve market power and monopoly advantages. I suggest below that the antitrust laws themselves may have an ambiguous affect on the size of the firm. See infra note 46. In any event, antitrust law is but tangentially related to the arguments in this Article.
mechanisms in place to attract principals’ capital at lower cost in the first place.\textsuperscript{13} From the perspective of the nature and size of the firm, it seems plausible, if not likely, that the superficial view is thus either wrong, or but one possibility. Firms may be expected to be smaller than a pure transaction costs approach suggests, because the fear of managerial tendencies and irreversibility will cause far-sighted principals to encourage their firms to be smaller than first seems optimal. Knowing that if their firm is too large it will be very difficult to contract it, principals may prefer that the firm remain small. This is especially the case if exogenous events, including business-cycle patterns and technological changes (affecting transaction costs), make the optimal size of the firm rise and fall. It will be unwise to let the firm (be it an automobile manufacturer, a law firm, or any other example that comes quickly to mind) grow to meet new profitable opportunities if such growth will lead to larger losses when circumstances change and contraction (to a point behind the earlier launching point) is the profit-maximizing strategy.

My argument might stand here. It suggests that the balance set out by Coase needs to be revised to reflect what I have called irreversibility. In particular, it is likely that profitable firms that pool capital and, therefore, face agency problems will substitute external arrangements (or no activity) for internal growth, because of the extra costs imposed by irreversibility. These may be the costs of aggressive monitoring, of forcing contractions on disinclined agents, of life as an inefficiently large firm, or of some other obstacles to contraction discussed presently. Some of these costs might be regarded as examples of transaction costs, in which case Coase’s sketch might simply be expanded to include the transaction costs associated with the stickiness I have described.

However, my aim in this Article is not only to complicate our perception of the balance between internal and external arrangements, and the determinants of firm size with the problem (or likelihood) of irreversibility, but also to explore the ways in which law causes or mitigates this problem. I suspect that every reader can think of ways in which law appears to raise the costs of contraction. I consider some examples in Part III.\textsuperscript{14} But the more difficult and surely more provocative question is: How might law reduce, rather than contribute to, irreversibility? My suggestion is that law is capable of such a role, and that it indeed occasionally rises to this task. It is also possible that collective action and other coordination problems have thus far prevented law from playing a more important role in reducing the costs of irreversibility. The central idea insofar as corporate law is concerned is that shareholders (or, more generally, principals) or agents seeking to bond themselves can combat irreversibility by installing appropriate charter provisions, off-the-rack default rules, or mandatory legislative rules. A variety of legislative and constitutional provisions, such as the Fifth Amendment’s promise of compensation for property taken for public use, or state provisions requiring balanced budgets or limiting tax increases, can be understood as precommitments intended to force


\textsuperscript{14} See \textit{infra} text accompanying notes 44-46.
better decision-making at a later time. Other legislation might anticipate the problem of irreversibility and constrain agents' power to expand their organizations' size and scope.

B. Mitigating the Irreversibility Problem

1. The Historical Example of the Ultra Vires Doctrine

One historical application of the theory suggested here concerns the rules governing actions engaged in by a corporation that go beyond the powers given to the corporation by its charter or the governing statute. A modern perspective of this ultra vires problem is that limited powers (such as a statement in the articles of incorporation that a corporation will engage in the railroad business, or a statutory provision barring corporations from owning stock in other corporations) can be understood either as a means of promising creditors that the firm will not engage in risk alteration after the terms of borrowing are set, or as a way of controlling the apparent authority of agents to engage in transactions that their principals would wish to veto. Limitations on the powers of firms offered principals a means to control managerialism. Agents may selfishly wish to expand the railway operations they manage into different states and into vaguely complementary lines of business. Principals can control such irreversible expansion by providing at the outset that the purpose of the corporation is only to engage in railroad operations in State Z. It is interesting to note that among the activities typically specified as beyond the powers of corporations was the ownership of stock in another corporation. Such ownership might have served to increase the reach of an agent's decision-making, and it could be accomplished without the agent's intermediation by principals' choosing to purchase stock in such other entities in direct fashion.

The ultra vires example is interesting only as a matter of theory and history. Modern corporations are generally established to make money in whatever way possible. Articles of incorporation reflect this general aim, as do the enabling state corporation statutes. The explanation for the disappearance of the monitoring provided by the ultra vires idea is probably that it was too strong. Principals may want to control their agents, but principals like to be ready and able to take advantage of profitable opportunities should they arise. Moreover, most creditors learn to constrain the firm's activities managers with contractual tools that imitate what the ultra vires doctrine had accomplished. This explanation works as well for each of the explanations of the old legal regime. The claim is simply that it is possible that

15. See generally ROBERT C. CLARK, CORPORATE LAW 675 (1986).
17. See infra text accompanying note 36.
18. FRANK H. EASTERBROOK & DANIEL R. FISCHER, THE ECONOMIC STRUCTURE OF CORPORATE LAW 102-03 (1991) (arguing that since one cannot tell in advance that a category of expenses, or presumably of investments, is always antithetical to investors' interests, courts now leave questions previously decided under the ultra vires doctrine to markets).
the law surrounding the setting forth of explicit and limited corporate purposes may have served to combat irreversibility—and the disappearance of this law may simply indicate that it was too costly a defensive weapon.

2. Obstacles to Corporate Expansion

A more powerful and ageless example of the relationship between irreversibility and the law, in which law may mitigate rather than exacerbate the problem, concerns the constraints on various means of corporate expansion. One focal point of this example is the asymmetry between acquisitions and divestments. If Corporation X wishes to acquire Corporation Y, X will generally need to jump through a variety of legal hoops to expand in this manner. For example, a merger or consolidation will generally require separate votes of approval by the two corporations' boards of directors and shareholders. If, however, after some time, X's managers decide that the acquisition of Y was unprofitable or otherwise unwise, X can sell off its Y division with no voting at all. This asymmetry provides the clearest example of the positive relationship between law and irreversibility. We might say that because X's managers are often too quick to pursue internal growth, their principals have established, or have invested under the protective umbrella provided by, institutions to check such growth. More to the point, the psychological and financial implications for managers from expansion are so different from those associated with contraction that principals put much greater faith in managers' recommendations to contract. If X's managers believe it wise to shed Y, X's principals may require no say in the matter, because, inasmuch as the decision goes against the usual managerial tendencies, it is quite likely that the decision is in the principals' (and society's) interest.

There are, of course, other ways to acquire and dispose of new lines of business or divisions. The theory suggested here becomes more appealing as other avenues of expansion and contraction are accorded legal treatment consistent with the dichotomy suggested above. X might engage in serious internal growth not by merging with Y, but by pursuing a number of other options. These options include the possibility that X issue new stock (and finance internal growth with the capital that is raised through this issue), buy assets from Y or any other seller, make a tender offer for Y's (or some other) stock, retain earnings and reinvest in X's growth, issue bonds or other instruments to raise investment capital, or simply borrow money from banks and other creditors. Most of these alternatives involve obstacles or antidotes to irreversibility, as we might optimistically call them, that are unmatched when comparable contractions are concerned. Sometimes expansion and contraction decisions must overcome similar, that is symmetrical, legal obsta-


20. Generally, no statutory provision governs something less than a sale of “all or substantially all” of a corporation's assets. Such sales may be approved by the corporation's board or an authorized officer in the same manner as any other transaction. See, e.g., Revised Model Business Corp. Act § 12.01 official cmt. 2 (1984). The point in the text does not, of course, depend on the previous (hypothetical) acquisition of Y by X.
Irreversibility and the Law

cles, but it is virtually never easier to expand than to contract. The broad thesis is that managers' lopsided inclinations to expand but not to contract are somewhat balanced, if not remarkably balanced, by legal rules that, at least in the aggregate, asymmetrically burden expansion more than contraction. A corporation with authorized but unissued stock can normally sell the previously unissued stock after overcoming the same obstacle, a vote of the board of directors, that must be met when repurchasing outstanding stock (and therefore contracting).\(^{21}\) But it is more difficult to sell truly new stock. The articles of incorporation will need to be amended through a shareholder vote, unless they originally authorized sufficient shares for present intentions,\(^{22}\) and, in any event, the securities laws require more of the issuer than of the repurchaser.\(^{23}\) Indeed, the very expense of complying with the securities laws may be thought of as an antidote to irreversibility.

Similar asymmetries exist where X buys Y's assets, or makes a tender offer for Y's stock. Generally speaking, if Y is small compared to X, there are few hurdles for X to overcome. This is as we would expect under any theory of these aspects of corporate law, because X's principals will want to delegate every day decisions if only to conserve transaction costs. But when Y is of substantial size, X's shareholders must vote their approval of an asset purchase, and, at a minimum, X will need to make disclosures to Y if it makes a tender offer to Y's shareholders.\(^{24}\) Again, one can hardly claim that these voting or disclosure requirements often provide substantive relief for principals (of X) concerned about managerialism and irreversibility. The idea is instead to think of these hurdles as costly, and, therefore, as slowing down managers' inclinations. And the interesting thing of course is the otherwise inexplicable asymmetry. There is virtually no legal hurdle to clear either when X's managers are ready to sell off division Y, or when X's managers tire of owning Y stock (acquired perhaps in a tender offer at an earlier date) and decide to sell this stock.\(^{25}\) The conventional wisdom about these hurdles to large asset or stock


23. The repurchase-new issue distinction follows from the 1933 Act's focus on the initial distribution of securities. Repurchases are of much less interest to the Securities and Exchange Commission unless they are part of a "going private" transaction. It should be noted, however, that some legal systems look much less favorably on—and even forbid—stock repurchases. I think it fair, however, to explain such a posture not as an example of an attempt to discourage contraction, but rather as a means of avoiding fiduciary breaches in the form of exploitative repurchases based on inside information.

24. The applicability of typical state law rules to mergers between entities of equal and unequal size is discussed in Clark, supra note 15, at 401-06, 414-18. Clark also provides an interesting comparison with tender offers. Id. at 546-54.

25. Tax law, however, discourages contractions of this kind. It is easier to satisfy the requirements of an acquisitive tax-free reorganization than it is to meet those of a divisive reorganization. But this asymmetry is conventionally (and correctly, I think) explained as aiming to prevent the bail-out of earnings and profits—that is, the escape of dividend taxation—through investment in a new division or subsidiary. See Howard E. Abrams & Richard L. Doernberg, Federal Corporate Taxation 1, 224-25 (2d ed. 1990).

A more sophisticated view of the tax treatment of divisions recognizes an asymmetry between the treatment of acquisitions and divisions, or contractions, and sees a complicated relationship between this
acquisitions is that shareholders are collectively experiencing a fundamental drastic change. But there is an equivalent change when a line of business is sold off, so that the absence of voting and other rights in that context suggests that there is something else at issue.

There is, it should be noted, the requirement of a shareholder vote when X wants to sell "substantially all" its own assets. Similarly, the dissolution of X requires a variety of formalities. In these cases of contraction, if they can be called that, there is, however, something else at stake. Students of corporate law are familiar with the conflict of interest that is often present when a target company's managers are offered employment opportunities in an acquiring company. It is not surprising when corporate law affords extra protections in situations in which a combination may take place at something less than arm's length. Here, when a corporation sells all its assets or dissolves (a procedure which often follows a substantial sale), managers are not generally contracting the firm so much as they are closing down an entity in what may be part of a transaction in which a new entity takes over. In any event, managers do not continue to manage what has become a smaller entity, and they may in fact control or be employed by the acquiring or surviving entity. The irreversibility idea suggests only that (society at large and) principals are relatively sanguine about managerial decisions that involve shrinking the empires these managers control. When an empire is not contracted, but rather sold off or reformulated, the irreversibility problem does not suggest a lower level of monitoring; indeed, if one takes into account the danger of side deals between acquirers and target managers, these may be appropriate occasions for heightened scrutiny. The law regarding dissolutions and complete asset sales appears to reflect this view.

We come now to the most interesting alternative means of expansion: debt and the retention of earnings. I will not dwell on expansion through earnings retention because there is a vast literature on this subject. From a tax perspective, it is sometimes puzzling that corporations pay dividends, and in this way voluntarily subscribe to the most raw form of two-tier taxation; the corporation has been taxed asymmetry, or non-neutrality, and that between the treatment of incumbent managers and raiders. See Paul B. Stephan III, Disaggregation and Subchapter C: Rethinking Corporate Tax Reform, 76 VA. L. REV. 655 (1990).

26. ABRAMS & DOERNBERG, supra note 25, at 403-04.

27. See, e.g., DEL. CODE ANN. tit. 8, § 271 (1988); MELVIN A. EISENBERG, THE STRUCTURE OF THE CORPORATION 255-62 (1976) (arguing that law should, but does not, provide for safeguards when there is a sale of less than substantially all the assets in a transaction not in the ordinary course of business).


29. A good example is the presence of the de facto merger doctrine, when managers of the target company and of the acquiring company may have come together too closely. See Hideki Kanda & Saul Levmore, The Appraisal Remedy and the Goals of Corporate Law, 32 UCLA L. REV. 429, 464 n.114 (1985) (linking case-law distinctions to presence of independent versus conflicted management, and arguing that the appraisal remedy itself ought to, and sometimes does, do the same).

on its earnings, and shareholders are then taxed on the dividends they receive out of the corporation’s earnings and profits, with no corresponding corporate deduction. Moreover, one would think that managers have every incentive to exaggerate the profitability of future internal investments to justify the nonpayment of dividends, the retention of earnings, and the internal growth of the firm. The most optimistic thing to say is that firms choose instead to raise capital through the issue of debt and preferred stock, and perhaps even to develop a culture in which dividends “must” be paid, precisely to combat the irreversibility problem. Readers of the corporate finance literature will recognize this possibility as a close relative of Jensen’s ingenious (and equally optimistic) cash-flow thesis, in which leveraged buyouts are understood as a means of forcing managers to the edge when they must produce enough revenue to pay debtholders and to survive. There is also something in common with Easterbrook’s view that dividend payments force firms more often to the capital markets, where the discipline imposed is valued by the shareholders who receive (taxable) dividends. I would hardly suggest that these theories, modified or even improved a bit by the idea that the cost of paying dividends may be partly offset by the value of the obstacle such distributions present to expansion-minded managers, completely explain the famous puzzle of why firms pay dividends at all. There are, after all, cheaper ways to contract. But I do suggest that the absence of legal hurdles to the payment of dividends is not by itself counter-evidence to the proposition that such hurdles are put up in disproportionate fashion in front of expansionary (as opposed to contractionary) steps. Moreover, there is the problem of identifying the proper starting point for evaluating the (absence of) legal hurdles surrounding dividends. We might note, for example, that it is far easier for the firm to pay dividends (to contract) than it is for the firm to assess its shareholders for new contributions (to expand). This asymmetry fits quite precisely the idea of law as a solution to the irreversibility problem. It goes without saying, however, that a better explanation for why we would not regularly expect or even permit such “anti-dividends” (or forced contributions) is that they could be used to put pressure on cash-poor or otherwise disinclined shareholders to exit from the enterprise. Put differently, managers who wish to expand must either convince shareholders to take a step such as issuing of new stock—so that a majority will generally either have veto power over an expansion, or be protected by preemptive rights—or must convince external lenders of the wisdom of the expansion. It is to

32. See Michael C. Jensen, Eclipse of the Public Corporation, HARV. BUS. REV., Sept.-Oct. 1989, at 61; Michael C. Jensen, Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, 76 AM. ECON. REV. PAPERS & PROC. 323 (1986). Note also that an optimist might say that government deficits are good because they force the government to try to raise taxes and thus to argue for each new project.
34. I do not mean to imply that common-law preemptive rights are terribly important. The simplest possibility is that if a majority of shareholders opposes an expansion, managers will be unable to issue new stock (unless it has previously been authorized). But even when an issue requires only the approval of the board of directors, it is possible that shareholders will object because of the proposed price of the new issue, or because managers will increase their control through the issue, and courts may
this external method of financing internal growth that I now turn.

When expansion is financed through a bond issue, the theme advanced here could be said to carry on. The expansionary step will involve costs imposed by securities law, but the firm can always enter the after-market and purchase its own bonds with no comparable hurdles to this form of contraction. The law can thus be said to counteract, or at least to constrain, managers’ inefficient (and expensive-to-reverse) propensities by imposing greater hurdles to expansion. On the other hand, this is a fairly weak asymmetry, because there is generally no requirement of shareholder approval for the new bond issue itself. Managers can expand through this form of leverage more easily than they can increase outstanding equity.

But there is no need to dwell on this form of expansion, because a more obvious method for managers to avoid legal hurdles to expansion is to approach a commercial lender, who might quickly arrange a loan to the firm. Statutes constrain dividends in the shadow of debt, and there are the now-familiar hurdles that must be cleared to expand through mergers and the like, but statutes do not constrain debt itself. The governing documents of a firm rarely place hurdles in the way of managers obtaining loans on the firm’s behalf. Balanced-budget amendments are the concern of those who monitor political agents in state capitals rather than corporate agents in Wilmington and New York.

The key here is to return to Coase’s central insight, and to revisit the link between The Nature of the Firm and law as a solution to the irreversibility problem. Internal growth and arrangements, it will be recalled, offer managerial flexibility, while external contracts, in return for higher transaction costs, offer the informational advantages of explicit prices. The irreversibility problem, in turn, is that managers may selfishly substitute internal for external arrangements (or for no growth at all). When the matter is put this way, it should become clear that, from the shareholders’ perspective, borrowing from a commercial lender is essentially an external arrangement. The lender cares not for an agent’s incentive or inclination to expand, but rather about the likelihood that interest and principal will be paid. In turn, the manager who wishes to expand must face the explicit price associated with the interest rate on borrowed money, and this price offers the advantages Coase associates with external arrangements. It is true that in the case of this external arrangement, the empire-building manager can use the arrangement to sponsor internal growth (so that the manager may not be much dissuaded by high prices for credit any more than the managers would pause if acquisitive mergers seem expensive), but insofar as the irreversibility problem goes, the lender is a fairly good monitor on behalf of shareholders. The bulk of the shareholders’ concern must be that their agents will expand the firm when the environment is profitable, but that

rescue these shareholders through some application of the doctrine of preemptive rights.

35. There is, of course, the problem that the lender’s agents will build empires by loaning when they should not. One response to this problem is that two sets of principals with agency problems may provide more monitoring than one. Another response hints at the comparative agency cost perspective that lies ahead. See infra text accompanying note 48. Creditors may be able to expand by lending to any of a number of potential borrowers, so that competition among the borrowers helps the lender’s principals, because the lender’s agents will prefer to expand more rather than less profitably.
they will be too slow or unwilling to contract when shrinking is the profit-maximizing strategy. This tracks closely the concerns of creditors; debtors will enjoy the upside return after paying fixed credit costs in cushy times, but will be unable to make payments during downturns. In short, the most important method of expansion that is free of legal hurdles brings on a monitor whose concerns are remarkably similar to shareholders' irreversibility fears. The contractual conditions imposed by this monitor are therefore likely to solve the irreversibility problem at least as well as statutes and other precommitments solve the problem through the erection of hurdles in the way of mergers, stock issues, and other expansive maneuvers.

3. Delaware's Anti-Takeover Statute

The preceding argument has suggested that the structure of corporate law, which often erects higher hurdles in the way of mergers and other expansionary steps than it does in the way of contractions or certain external arrangements, supports the idea that law can anticipate and mitigate the irreversibility problem. There are less systemic legal rules that might be similarly understood. Consider, for example, Delaware's anti-takeover statute. It is a "moratorium statute," prohibiting "business combinations" with an "interested shareholder" for a three-year period following the acquisition by any shareholder of more than fifteen percent of the voting stock, unless the board of directors has pre-approved the acquisition. The prohibition, or moratorium, does not apply if the acquirer buys more than eighty-five percent of the stock (presumably because this larger acquisition leaves fewer shares to receive less in a lower-priced second step, and therefore minimizes the potential prisoners' dilemma problem faced by target shareholders), or if, subsequent to the acquisition, the board of directors and two-thirds of the other shareholders approve the business combination. The interesting thing, however, about the statute from the perspective of the irreversibility problem is that, unlike other moratorium statutes, the prohibited business combinations include only those between the target and the acquirer and its affiliates. Most prominently, the statute does not affect an acquirer whose plan is to disassemble, or "bust up," a target and distribute the proceeds to the remaining shareholders as a large-scale, dissolving dividend. As a matter of interest-group or interstate politics, this seems quite surprising, because anti-takeover statutes are often understood as inefficient attempts by managers and states to maintain what they enjoy; such a view suggests that bust-ups would be at the top of the list of transactions to be stopped.

36. Similarly, we might say that the public issue of debt also offers a market check on expansion, because of the explicit signals and costs in that market. Securities law constraints on public debt can be seen as ensuring information for creditors, and also that the market's signals are meaningful and not the product of some quick and misleading scheme.

38. Some comparisons among the statutes are offered in Choper et al., supra note 16, at 1094-98.

40. See Choper et al., supra note 16, at 1097.
41. Many other business combinations leave the target in place so that managers continue in their
In contrast, the irreversibility theme explains this apparent oddity quite neatly. A free and well-functioning takeover market is thought, perhaps too hypothetically, to discipline managers, because inferior managers will be displaced through the takeover process. But to the extent that the problem is that managers selfishly over-expand their firms, it is something of a solution for better managers to take over the firm—because better managers are, in Coase's terms, best matched with larger internal organizations—but it seems far more likely that the efficient solution is to contract the firm. Disassembly is thus the most powerful tool ex post, and even ex ante it is likely to be an excellent threat to discourage over-expansion. The Delaware statute's remarkable exception for bust-ups can in this way be understood as a terrific (and efficient) example of law as the provider of antidotes to the irreversibility problem.

It goes almost without saying that we had better not make too much of this example. It seems unlikely that the Delaware moratorium statute as a whole is good for shareholders and the takeover market, and only a strange evolutionary story could therefore account for shareholders, managers, or public-regarding legislators doing great damage with their broad brushes, but, meanwhile, cleverly precommitting in an efficient manner in one small piece of the larger devastating picture. It is not an impossible story. The dominant forces and interest groups in Delaware may have succeeded in drafting a statute that copied New York's more restrictive moratorium-style anti-takeover statute, and the irreversibility problem may have caused a loosening of the statute exactly where we would most have wished it. But many other states' statutes also make no exception for bust-ups, and given the interest groups that appear to have captured the legislatures that pass these anti-takeover statutes, it will hardly be surprising to find statutes putting the greatest hurdles in the way of disassemblies. In short, the Delaware statute offers an oddity that may well be explained in the suggested manner. It is the broader argument, however, comparing the relative hurdles in the way of merging of two companies, casting off a division, and borrowing from a commercial lender (to expand) that makes the positive case for law as a partial solution to the problem of over-expansion and irreversibility.

III. AGENCY COSTS AND LEGAL COSTS

A. Law as a Source of Transaction Costs and Irreversibility

The discussion in Part II suggested that law may be a means for principals and agents, or for societies in general, to constrain the private incentive of agents to
overexpand their firms. In turn, the basic structure of corporate law, or at least of the rules governing organic corporate changes, suggests that this possibility may form the core of a positive, descriptive theory. This set of examples, however, has surely not taken too much away from the more familiar role law—and perhaps especially law aimed at large businesses—plays as a preserver and protector of expectations associated with the status quo. In this role, law sometimes constrains expansion because growth threatens existing values and interest groups. It is far more common, however, for law to protect interests by standing in the way of contraction. Law can in this straightforward way be a cause of, rather than a solution to, irreversibility. When this is the case, the optimal size of the firm in a dynamic setting is presumably smaller than that suggested by The Nature of the Firm, because expansion can often be undone only at the substantial cost dictated by protective legal rules.

It may be useful to restate this point to avoid confusing the legal rules referred to presently with those discussed in the previous Part. The theme of Part II was that some legal rules can be seen as off-the-rack, even precommitment-style, solutions to the irreversibility problem. Individual agents may tend to overexpand, but law may counteract this tendency by making certain kinds of expansion more expensive than contraction. There is, of course, the more intuitive idea that law protects vested interests, or, more optimistically, investment and reliance interests. Consequently, in this role, law may make contraction more expensive than expansion. The most natural way to think of these conflicting roles is to think of the sources of law. Pieces of corporate law, for example, can be seen as combatting irreversibility, while plant-closing laws, for example, can be seen as causing the irreversibility problem. In terms of the sources of these laws, shareholders and (precommitting or bonding) managers may influence the substance of corporate law, while other interest groups may bring about the enactment of such measures as plant-closing laws. Environmental and labor laws are influenced by still different interest groups. Positive political theory in general, and the structure of legislative committees and other institutions in particular, lead us to expect that given interests will gain in some arenas even while experiencing setbacks in others. Much as heavy manufacturers might, as an interest group, experience a mix of new environmental, tax, labor, and tort laws in a given time period, such that it would be difficult to say whether in the aggregate this group was better or worse off than before the new decrees, so too law may at the same time, but with assorted provisions, generate and mitigate the irreversibility problem. And much as we would miss an important part of the story if we ignored the interest that politically active manufacturers might have in lower taxes and specific safety regulation, so too we would misunderstand other areas of law if we did not consider the role of law as an obstacle to over-expansion.

44. The protection of these costs can, of course, be efficiency enhancing. Thus, plant-closing laws may force a firm to internalize the costs workers face in securing new employment. Workers will want to begin searching and re-training before separation, and plant-closing laws encourage firms to give advance notice (or to pay and internalize some of the costs of surprise). See Workers' Adjustment, Retraining and Notification Act, 29 U.S.C. § 2102 (1988) (requiring certain large employers to provide 60 days advance notice of intent to close a plant).
The view just sketched assumes that there are at least some areas of law in which rules make contractions more costly than expansions. This straightforward asymmetry is familiar to most lawyers in the form of the sometimes ridiculed and deconstructed, but always appreciated, distinctions in so many fields of private and public law between action and inaction, regulatory constraints and subsidies, and so forth. It is not only psychologically easier to hire than to fire, it is also legally easier to convey good news and new jobs. The entrepreneur who hires new workers when expanding faces some legal constraints that might conceivably discourage expansion (and encourage external arrangements)—most notably, potential lawsuits by persons passed over in the hiring process. But discharging workers as part of a contraction runs far greater legal risks. Suits for wrongful discharge are far more likely than those for wrongful non-hiring. Although there are many reasons for this differential, we might simply note that the identity of persons injured by discharge is much clearer than the identity of those injured when the "wrong" persons are hired. This difference translates, as it does in many areas of law, into an important variation in the likely success of suit.

Although there are many examples of this asymmetry (such that law causes irreversibility), there are, of course, contrary examples, or even fields of law. A firm that wishes to shut down a plant may face union difficulties, and even plant-closing laws, but one suspects that these costs are overwhelmed by those tossed up by other fields of law—or legislative committees responsive to other interests—in the face of a firm that decides to open a new plant. Unless this expanding firm finds itself wooed to an industrial park, or to some other location or line of business where many regulatory hurdles have been cleared at the outset of development, there are likely to be substantial transaction costs associated with zoning issues, building permits, safety regulation, and environmental controls. Moreover, inasmuch as some of these legal hurdles are best understood as supported by interests that fear business expansion, not (only) because of the fear that the plant costs imposed will be inappropriately externalized and ignored, but because of a historical ideal glorifying farms and corner stores rather than corporate headquarters and discount chains, it is appropriate to think of these legal costs of expansion as comparable in some sense to those that stand in the way of contraction. There is a sense in which all these rules are enacted to preserve the status quo.

I do not mean or need to assess the precise magnitude of the legal obstacles to contraction and to expansion. Indeed, there are areas of law, such as antitrust, in which it is difficult to assess the net effect of even a single statute on firm size. My


46. Even when antitrust laws are taken very seriously so that internal growth is discouraged because of the potential problem of size and monopolization claims, other aspects of these laws must encourage internal growth by discouraging external arrangements—through the obstruction of vertical restrictions, for example. I do not emphasize the role of antitrust law as an antidote to managerialism in this Article also because of the ambiguous impact of this law on the irreversibility problem. There is much in the law's origin and remedies to suggest that principals would prefer to do without the constraints on their firms that this law imposes.
argument is mostly about the positive interaction between law and the irreversibility problem, because that is the face of law that I think has been unseen. My own sense, however, is that there is not an insignificant "negative," or causal, interaction between law and irreversibility, with law often preserving the status quo by blocking contractions. If so, then legal rules may be a cause of irreversibility, and, in turn, principals may wish to expand more slowly than intuition first suggests, because they know that legal rules will make subsequent reversals expensive to accomplish. If my sense of the legal landscape is controversial, and other observers guess that environmental and other anti-growth regulation inhibit expansion so much more than they do contraction that there is little need to think of law as the cause of irreversibility, then little is lost in thinking that irreversibility is caused by managerial incentives and psychology.

B. Reinterpreting The Nature of the Firm: Relative Agency Costs

The argument up to this point, and The Nature of the Firm itself, slides over an opportunity-cost criticism that is usefully raised at this point. Coase (and the discussion here) describes the firm as balancing internal flexibility—that is, relatively low (internal) transaction costs—with the advantages of signals from external arrangements. The firm's size is thus a function of an internal calculation, or optimization strategy. The analysis is of the standard microeconomic variety, in which one firm's profit-maximizing decision-making is analyzed in a marginalist way.

The problem with this view is best seen by considering a firm, X, that chooses to externalize an arrangement, perhaps by contracting to purchase an input from an external maker, Y, and perhaps for the reasons set forth by Coase. What about the nature and size of this firm, Y? The arrangement that X has made external is now internal to Y (or to some other firm that Y contracts with), so that the cost-benefit calculation made by X in deciding whether to do something internally or externally is now mirrored in Y. The transaction costs of internal growth cannot really be avoided by X's deciding to externalize an arrangement, for they are now borne by Y which, in turn, charges X for these costs (plus a premium, we might even presume) as part of the contract price charged X. If the interesting question is how an economy will be divided into firms—for example, when there will be several large firms or many more smaller firms—then the answer must focus not on how a given firm idealizes internal flexibility and external signals, but rather on how a given firm's agency costs (or terms of tradeoff between internal flexibility and external signals) compare with other firms' costs. It will be profitable for X to forsake internal growth in favor of an external arrangement with Y if Y can absorb the costs of growth more cheaply than X, or if X needs external signals more than Y (or more than another firm that Y in turn contracts with in order to receive an external signal itself). X will form an external arrangement with Y if these comparative advantages are serious enough to warrant the transaction cost of the explicit X-Y

---

47. As opposed to the question of when transaction costs will simply stifle growth.
contract. Growth in the economy will be allocated among existing and new firms based largely on their relative agency costs.

This relative-agency, or opportunity-cost, view of the nature of the firm reminds us of the omnipresence of agency costs in understanding economic arrangements. Managerialism and aggregate demand aside, firms will grow if they are able to do so more cheaply than their competitors. This cost comparison is in large part a question of efficient supervision and organization, or perhaps a function of specialization (in various monitoring tasks) by managers of different firms. X will externalize its arrangements when other firms can produce things and transfer them to X more cheaply than X can produce on its own. It is thus the relative agency costs (plus a transaction cost) that separate X and Y. In turn, this opportunity-cost view casts a different light on law as a cause of irreversibility.

Reconsider, for instance, plant-closing laws as an example of a legal rule which may make contraction difficult, and, therefore, contribute to the irreversibility problem. It should now be clear that this example is misleading. It is surely correct to say that plant-closing laws make contraction more expensive, and might therefore cause far-sighted investors to "underexpand" and the economy as a whole to suffer (either because contraction is too expensive, or because there is resulting underinvestment). But if the aim of *The Nature of the Firm* is to understand the division of economic activity among firms, and to understand the size and scope of individual entities, then the question is whether a firm can really avoid the potential cost of a plant-closing law by substituting an external arrangement for internal growth. In the case of state law, the migration of activity to a less regulated jurisdiction is a familiar possibility. However, in the case of federal law, why should the cost of the law be any lower if one firm externalizes an arrangement, and thus shifts it to another firm that is also subject to the law? The apparent conclusion is that such costs do not on their own contribute to irreversibility, or to our earlier understanding of the nature of the firm. If there is any microeconomic effect at all (that is something other than less contraction or, in the long run, investment in general), it is probably the transfer of activities to firms that are either less likely to close plants when this activity or others prove unprofitable (perhaps because they have diversified, profitable opportunities appropriate to a given plant) or unregulated by the plant-closing legislation, perhaps because they are too small to be covered by the statute.

This substitution effect may be of fairly general application. If, for example, employment law contributes to irreversibility, because lawsuits brought under the anti-discrimination statutes are more likely or of greater expected cost when firing rather than hiring is involved, then activities may well be transferred to, or even subdivided among, small firms with too few employees to be governed by these statutes. These avenues of escape suggest the rather obvious point that the size and

---

48. There is a hint of this in Ronald H. Coase, *The Nature of the Firm: Influence*, 4 J.L. ECON. & ORGANIZATION 33, 39 (1988), although the focus appears to be on other firms' operating costs, rather than transaction, or agency costs.
49. 29 U.S.C. § 2102 (1989) (covering employers with 100 or more full-time employees).
50. Title VII, for example, applies to firms with more than 15 employees, 42 U.S.C. § 2000e(b)
Irreversibility and the Law

scope of the firm may have something to do with cost minimization in the shadow of the law rather than with the sorts of considerations emphasized by Coase or in this Article. Put more negatively, while it is easy to think of many laws that seem to contribute to irreversibility by threatening to impose compliance and other costs on firms that seek to withdraw from an activity or internal employment relationship, when these costs would not be imposed on a firm that simply failed to renew a contract with an outside supplier, virtually all of these laws apply just as well to the outside supplier that our firm might turn to in lieu of internal growth. The costs of these laws are avoided only by crossing jurisdictional boundaries, by generating less activity in the first place (rather than by engaging in internal-external balancing), or by substituting in favor of entities that are not subject to the laws either because of their small size, not-for-profit status, or other special station. Moreover, such substitutions—that is, potentially inefficient biases in favor of entities that are untouched by, or otherwise favored by, law and regulators—are familiar, and hardly require the framework offered by The Nature of the Firm or the irreversibility idea. Put differently, much of law may be about the preservation of status quo interests, but these laws and the political influences behind them tell us little about the size and scope of a firm, because they generally apply not only to internal but also, however indirectly, to external arrangements.

From the perspective of internal-external decisions, the most important example of law's favoring the status quo may be the tax law's pattern of awaiting recognition events before accounting for gains and losses. Tax law might be said to contribute to irreversibility to the extent that a background rate of inflation, or the law's own system for allowing depreciation expenses, causes most assets to appreciate in value over time, at least compared to their adjusted bases. Tax law thus discourages contraction, because the sale or other disposition of assets triggers gain recognition. In contrast, the acquisition of assets often triggers no gain recognition either because the acquisition is financed by internally retained earnings, in which case distribution would trigger dividend taxation in the case of corporations or no further recognition in the case of partnerships or similarly taxed entities, or because the acquisition is accomplished by new debt or equity, in which case, again, the structure of tax law generally favors the pooling of capital at least in the sense of deferring recognition.


51. It should be noted that Coase specifically allowed that firm size or existence would be a function of taxes and other government interventions. CoAsa, supra note 1, at 41.

52. In the period between acquisition and sale, many assets will have appreciated in value, perhaps because of inflation. Other assets may have depreciated, but their business owners will have enjoyed depreciation deductions, which in turn lower the "adjusted bases" of assets. Any excess of these deductions over real depreciation in market value as revealed at the time of sale is recaptured at the time of sale. See I.R.C. §§ 167, 1016, 1245, 1250 (1992). And given generous depreciation schedules, many assets will contain such potential tax liability (unless more than offset by the lemons problem in the used goods market) even though the assets themselves have been part of a losing project. Contraction can in this way be discouraged by tax law even when the assets are themselves unprofitable.

53. See Hideki Kanda & Saul Levmore, Taxes, Agency Costs, and the Price of Incorporation, 77
But this example of law asymmetrically favoring expansion over contraction again says very little about the choice, be it for a given firm or for all firms, between internal and external arrangements, except to say what every tax lawyer and devotee of capital gains preferences knows: that once a firm has invested in resources, it will hesitate to dispose of them to the extent that there is a tax bite upon disposition. In the short run, and once assets are already in hand, X may be locked in to the assets it possesses so that it may prefer to use these assets internally rather than dispose of them and contract with Y. But this is nothing more than the familiar lock-in effect of a tax system which awaits recognition events. In the long-run, this tax effect might either encourage X to expand, perhaps because the treatment of acquisition and depreciation is favorable in a way that is not easily replicated by Y and passed on to X in the form of cheaper external prices, or discourage X from investing internally, perhaps because the implicit taxation of inflationary gain upon later disposition or contraction discourages X from expanding or investing in the first place. But whether the net effect favors expansion or contraction, the same effect must normally pertain to Y and to other firms. And to the extent that firms may (for tax reasons) lease rather than purchase equipment, both X and Y can turn to this substitute. More generally, short-term leases may be a means of protecting against over-expansion, but X and Y face the same transaction costs in considering the price of this particular precommitment strategy. In short, tax law may affect the decision as to the form of investment, and it may affect the decision as to whether and when to sell or otherwise recognize gains or losses. But it would seem to have relatively little to do with whether X carries out a task through internal expansion, or by way of an external arrangement with (the now larger) Y. What matters is the relative agency costs encountered in X versus Y's expansion.

To be sure, tax law can be an important source of irreversibility (or of stimulus to expansion), because the law often treats different entities in dissimilar fashion. With the exceptions noted earlier, plant-closing laws generally treat X and Y the same. But an operation may be profitably located in Y rather than X if, to cite a few obvious examples, one is a corporation and the other a partnership or proprietorship (or the other way around), if one has net operating losses to use, or if one has other sources of income against which to offset the early losses that might be expected in the early phase of this new activity. Again, however, these differences have long been familiar to lawyers. Rather than making the case for the analysis proposed in *The Nature of the Firm*, or for that suggested by the irreversibility theme, lawyers suggest that in most instances in which law is really a source of the irreversibility problem or of the decision to enter into an external arrangement rather than to pursue internal growth, we need to focus more on either the distort-

---


54. Tax laws seem therefore to have a profound influence on a firm's mix of inputs, as capital and labor (and different kinds of capital) are subject to different treatments. The impact of these differences is not obvious; to the extent that labor costs are deductible as paid, for example, most of these payments (unlike many payments to capital investors) are immediately taxed as ordinary income to recipients. The effects of these laws are quite beyond the scope of this Article.
tions or non-neutralities arising out of tax law, or the relative agency costs encountered in the expansion of different firms. The balance a given firm's entrepreneur experiences between internal flexibility and external signals seems less important than before.

This comparative cost reinterpretation of *The Nature of the Firm* does not, it should be noted, detract from the positive interaction between law and the irreversibility problem explored in Part II. Principals will wish to offset the tendency of their selfish agents to favor expansion and to resist profitable contraction. In this regard, they may appreciate corporate law rules which place higher costs on expansion than on contraction. The fact that Y's principals do this even as X's do is perfectly consistent with the idea that X's managers will be biased in favor of internal growth over external arrangements and, in turn, these managers' principals will favor obstacles to internal growth. It may be that the effect of all this is to generate external arrangements involving firms with lower agency costs. This is as we should expect; agency costs of various kinds encourage the emergence of sole proprietors, for example, who have much less (and perhaps no) incentive to over-expand.55 The picture is one of firms that expand both to take advantage of economies of scale and, unfortunately, in response to managerial preferences. The firm's size is thus related to real economies of scale, to uncontrolled managerialism, to tax rules (and other factors), and to relative agency costs. Some entrepreneurs will simply be more efficient managers than others, and perhaps we should understand the argument in *The Nature of the Firm* simply to be that the more efficient the manager, the less the need for external price signals.

### C. Predicting Firm Size

The irreversibility problem, along with so many other varieties of agency costs, can be avoided by minimizing the divergence of interests between principal and agent. On the other hand, an economy consisting entirely of sole proprietorships does not avoid agency costs, although we may label them differently, because these costs simply reappear in the form of costs associated with external arrangements. These external arrangements may also be costly to form (although they may provide useful explicit price signals), but the point is that contracting parties will need to monitor each other's performance much as entrepreneurs need to monitor their employees and shareholders their agents. There may be a bit less to monitor externally, but when this occurs, it is generally because some of the transaction costs have been pulled forward to the time surrounding the formation of contracts. Still, it is interesting to note that at least the irreversibility problem is largely absent in a world of sole proprietors.

More generally, closely held firms can be seen as representing a form of organi-
zation that competes with that which generates publicly held firms. The former economize on some agency costs, including those associated with the irreversibility problem, and to the extent that closely held firms are smaller, they also enjoy more external arrangements and explicit signals. Large publicly held firms, in contrast, delegate more decision-making, suffer far more substantial free-riding problems because of their numerous and dispersed principals (and thus have more of an irreversibility problem, among other agency costs), and are likely to have proportionally fewer external signals to guide the allocation of inputs. Offsetting these disadvantages is not only the transaction-cost saving associated with fewer external arrangements, but also the “external” monitoring provided by public trading. A few well-informed monitors, by trading for their own self-interest, keep managers, other principals, and potential acquirers advised.66

One natural question this description raises is whether we can predict the size of firms or changes in the ratio of closely held to publicly traded firms. I do not pursue this question here because it is too daunting an empirical project. Any theoretical insights we have can be easily swamped by tax laws, technological changes, and demographic shifts, to name just a few variables that might encourage or discourage the organization of activity into firms of a given size. Still, if over time there were fewer small firms, one might surmise that the irreversibility problem—among other agency problems—was not much of a problem. It is therefore interesting to note that the formation of new firms appears to have out-paced the growth of gross domestic product.68 I do not advance this single fact as any sort of proof that over time firms not only externalize arrangements but do so in favor of small firms so as to reduce agency costs. Nevertheless, if the data showed the reverse it would give pause. More generally, it should be noted that the propositions put forth in The Nature of the Firm are themselves virtually impossible to test. What I have suggested is simply an extension to, or modification of, that way of thinking.

IV. THE IRREVOCABILITY PROBLEM IN OTHER ORGANIZATIONS

A. Other Asymmetries

Although this is an occasion to focus on the nature of business firms, the irreversibility theme is at least as interesting in other contexts. In virtually all

56. Potential acquirers can, of course, take over the firm and subdivide it so that more explicit signals are available to each segment. Such disassembly also resolves, temporarily, the over-expansion problem.

57. It is not obvious, however, which technological changes increase or decrease the tendency to internal growth over external arrangements. Lower transportation costs and facsimile machines, for example, lower the costs associated with external contracts, but they must also lower a manager’s costs in monitoring such things as multiple plants.

58. For suggestive data, see U.S. Bureau of the Census, Statistical Abstract of the United States 432, 532 (111th ed. 1991) (growth in number of business establishments greatly exceeds growth in the GNP). The enormous growth in the number of small firms is presumably correlated with the statistic that would be of greatest interest—the number of formal external arrangements. These arrangements, it should also be noted, might be well correlated with the volume of litigation.
organizations, we might expect agents to over-expand for a variety of self-serving reasons, so that the question is whether bonding, monitoring, or other mechanisms have been created to make expansion more difficult than contraction in the first place. A number of the most dramatic agency and irreversibility problems should be set aside at the outset, however, because on close inspection they do not resemble the phenomena emphasized here. Consider, for example, the possibility that generals thirst for victory and glory so that they attack and invade more quickly than is optimal. Well-informed citizens, and certainly infantry, might prefer defensive positions and retreats (or even peace) more than their military officers, and there is certainly a sense in which a failed attack is costly and thus irreversible. But because there are also generals who have been thought too cautious, perhaps because they perceived their self-interest as avoiding palpable setbacks, it seems most sensible to think of this agency problem in its most obvious terms; agents may have different risk preferences from their principals and this misalignment can be costly to principals. The jargon of irreversibility adds little to this more familiar perspective. Even if it could be shown that there is more delegation of military authority to retreat than to attack, the better explanation would seem to have more to do with the need for quick decision-making than with a far-sighted solution to the irreversibility (or risk-seeking) problem.

A better example is provided by the expansion and contraction of political jurisdictions. It is tempting to say that leaders may have too strong an incentive to push for growth during their terms. In the United States for example, the substantial hurdles to statehood offer a neat example of rules meant to offset agents’ private incentives. One problem with this view is that the hurdles to expansion are not greater than the legal obstacles to contraction; it is even more difficult and perhaps impossible for a state to secede. This comparison suggests that a better approach to this area is to think of the hurdles to contraction as necessary to prevent exploitative behavior by majorities threatening to expel members, and by minorities threatening to abandon the Union. It is easy to think of arguments in favor of a regime that simply does not permit secession. In turn, such a regime should be expected to place serious hurdles in the way of new entrants.

59. In the case of the United States, however, it is interesting that although we sometimes associate territorial acquisition with individuals, such as the Louisiana Purchase with Jefferson (and with perfect foresight, among his other talents, he might have had too strong an incentive to recommend the acquisition), we do not generally associate the incorporation of new states with individual agents.

60. Experimenting with expansion is especially sensible when it is easy to contract if the experiment proves unprofitable. Democracies, however, have reason to fear secession because of the problem of coercion by minorities threatening such action. In turn, because “independence” movements are unpalatable or expensive, expansion in the first place can be very expensive. This irreversibility idea may help explain the notion that democracies seem less inclined to war than despotistic states. Not only must democracies be more responsive to their citizens, but also they may find it more difficult to avoid sharing governing power and benefits with the “losers” (especially if these losers live in contiguous areas). Non-democratic forms of government may find it easier to extract gains from new residents, and even to do so on differentiated terms. In any event, firms may expand more readily than democracies (and perhaps even other sovereigns) because they can dispose of parts later.

61. The argument in the text might also be made in terms of “exit” and “voice.” Inasmuch as it is
An obvious analogue is the relationship between the size of hurdles to marriage and to divorce. Without expressing any view on the wisdom and morality of state or religion-imposed constraints on these agreements, it is safe to say that it is not irrational for a system of social control to bar divorce, or at least to put enormous hurdles in the path of most attempts to divorce. And once we imagine systems that make divorce impossible, or nearly so, we can predict that cultural or other norms may make entering into marriage a serious and even difficult proposition. This can be seen as the converse of the irreversibility pattern described earlier. In the corporate context, rules raise the cost of that which agents are thought to overvalue, namely expansion. In the family setting, “agents” may have a propensity to quit, or contract selfishly, when things are difficult. The rules can thus be understood as aiming to make contraction difficult, and there is the obvious resemblance to the irreversibility theme.

Within each political jurisdiction, there is also irreversibility of the kind experienced by firms. It is more difficult for the government to abandon, or even to privatize (and subsidize), activities than it is for the government to expand into these roles. Ideally, a government’s agents ought to be evaluated according to how well they perform with the resources at their disposal. Even when citizens “exit” by buying better security or schools in the private sector, it may be that the government is performing well. But our tendency to ask such questions as whether a locality has “good” schools or universities or playgrounds, as opposed to whether its institutions do well with the resources at their disposal, surely contributes to the kind of irreversibility caused by civil servants’ own incentives to expand.

B. The Irreversibility Theme in Non-Profit Organizations

In this article, I have concentrated on the irreversibility problem as it pertains to the corporate firm and its close competitors. In this context, I suspect that although many readers are skeptical of my optimistic view that law can be seen as offering partial solutions to the problems of managerialism and over-expansion, few readers disagree that there is a problem to be solved. There is enough in common experience, and in the literature on defensive tactics, executive compensation, and bureaucracies in support of this perception that an expression like “empire-building easier to exit from a firm than from a nation, voice is more important in nations than in firms. In turn, this may have an effect on optimum expansion.

The expansion (by annexation, for instance) and contraction of political units within a nation offers another example of the phenomena discussed in this Article. It is an example that I do not pursue except to note that expansion (and contraction) need not face as much of a hurdle, because there is an external monitor (such as the state, in the case of expansion by a municipality) in place.

62. Elizabeth S. Scott, Rational Decisionmaking about Marriage and Divorce, 76 VA. L. REV. 9 (1990) (discussing premarital contracts and other precommitment strategies that vary with the ease of divorce).

63. One might also describe decisions to delegate authority to administrative agencies as external arrangements, at least as compared to the internal growth of legislative involvement.

64. Thus, a fast-food franchise and an expensive restaurant are not regarded as failures when potential customers choose to patronize one rather than the other.
corporate officer" requires no explanation. Indeed, no opposite term comes to mind, because we do not think of many business agents as pathologically devoted to the leanness of their domains.

It is possible, however, that we overestimate the problem of over-expansion and underestimate the constraints and incentives offered by principals, the employment market (which may value a record of cost-cutting more than it values the scope of an applicant's previous responsibilities), capital markets (including securities markets and commercial lenders), and the threat of takeovers. I have, to be sure, neither direct evidence nor a means of testing the size of the irreversibility or over-expansion problem, but there is a piece of what we might think of as negative or comparative evidence. Voluntary corporate contractions (but not everyday decisions to substitute external arrangements for internal growth) may be more scarce than voluntary expansions, but the comparison pales in the presence of the non-profit sector. One can hardly locate examples of hospitals, universities, law schools, opera companies, or museums voluntarily contracting. There are instances of universities closing down departments or whole areas of instruction, but these are usually overwhelmed by roughly concurrent expansion on other fronts. Even when this is not the case, the ratio of expansionary to contractionary initiatives must be far greater for universities and most other not-for-profit organizations than it is for business organizations. It is difficult to name a law journal, law school, university, hospital, or museum with fewer editors, students, beds, or works of art than were present several years ago. Yet, surely there is reason to think that some of these organizations would be better, however their goals are defined, or that the society they serve would be better served, with some contraction. Put in terms of agency theory, few such organizations seem to maximize average quality (which might maximize the value of reputation or diplomas, for example). Yet, it seems likely that more specialization (as a result of some managers' limiting their scope to what they do best) would sometimes be efficient and perhaps in the interest of principals, however they be identified.

As already implied, one explanation for this apparently more serious irreversibility problem in the non-profit sector is simply the absence of any market check. The expansionary instincts of managers of corporations are more controlled than those of non-profits, the argument goes, because the former are monitored by stock markets or, in the case of closely held firms, self-interest (or self-interested co-owners).

More interesting, perhaps, is the possibility that evolutionary mechanisms provide greater controls in a way that is correlated with more serious irreversibility problems. It may be true that the great universities contract less often than the great corporations, but it is even more striking that the most successful universities expand less than their profit-oriented counterparts. It is common to find modest growth over time in student populations (even holding co-education and post-graduate trends aside) and in the number of fields of academic inquiry. But most

65. Both sectors are monitored by lenders, customers, and employment markets for the managers' skills.
opportunities for expansion, including franchising, multi-state operations, and tuition-revenue maximization are entirely avoided. There are explanations for this pattern that are unrelated to the irreversibility thesis advanced in this Article. One possibility is that managers and principals, be they alumni, students, taxpayers, or past and present donors, have inclinations that are quite well aligned, so that there is no agency problem to be solved. Both groups may seek to maximize status (such as rankings or public perceptions of the elite) so that expansion is undertaken cautiously and there is virtually nothing to be gained (on the part of agents, for example) by maximizing revenue or franchising. In this regard, it is interesting to note that while corporate managers' compensation is thought to be correlated with the size of firms, I know of no evidence suggesting that university presidents' compensation or status is correlated with organizational size. This approach does not, however, fully explain the failure of universities and similar organizations to contract more often.

A more concrete, but slightly different, version of this argument is that one aspect of the genius of the firm is that a successful firm's owners can capture the value of what they have created by selling new ownership interests in a way that generates profit for the pre-existing owners. This is much more difficult for non-profit organizations. A university that improves and innovates will have a difficult time raising tuition in a way that benefits the previous generation of principals, if only because its tax-exempt status forbids the distribution of profits to these principals, whoever they may be. In light of this inability to maximize profits, at least from the perspective of individual donors or other principals, it is especially unsurprising that profitable expansions are not undertaken. The principals may have little to gain and a great deal to lose.

In any event, I leave for another effort a more sustained inquiry into irreversibility in the non-business context. Such an inquiry will need to explore the impact of tax exemptions, governmental monitors, and other features that distinguish the non-profit sector. My aim here has been only to introduce the irreversibility idea, along with the companion notion that law (and other arrangements) can mitigate this agency problem. I think it is likely that the irreversibility problem is more serious in the not-for-profit sector, and I think it is interesting that there are parallel solutions in the form of cultural and grantor-initiated constraints on expansion, but my argument is meant to be more suggestive than insistent.

66. Data do suggest a role for profitability. See Scherer, supra note 12, at 36-37. But the point in the text is that while we expect the managers of the largest industrial corporations to be among the highest in the nation, we do not look for a list of the largest universities in order to guess the managers with the largest salaries in that sector.

67. Reputation does seem to be correlated with endowment size, but this connection may simply restate the question of how and why certain kinds of expansion occur. For an interesting discussion (that raises more questions than it resolves), see Henry Hansmann, Why Do Universities Have Endowments?, 19 J. LEGAL STUD. 3 (1990).