Neoclassicism and the Separation of Ownership and Control

Herbert J. Hovenkamp
University of Pennsylvania Law School

Follow this and additional works at: https://scholarship.law.upenn.edu/faculty_scholarship

Part of the Business Organizations Law Commons, Corporate Finance Commons, Economic History Commons, Economic Theory Commons, Industrial Organization Commons, Law and Economics Commons, Legal History Commons, and the Securities Law Commons

Repository Citation

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.
NEOCLASSICISM AND THE SEPARATION
OF OWNERSHIP AND CONTROL

Herbert Hovenkamp†

INTRODUCTION ...................................................................................................... 374
I. FISHER’S SEPARATION THEOREM ................................................................... 383
II. VALUE MAXIMIZATION AND THE NATURE OF THE FIRM ....................... 286
III. THE FINANCIAL STRUCTURE AND VALUE OF THE
    CORPORATION .................................................................................................. 391
    A. Modigliani-Miller: The Fungibility of Ownership and Debt................. 393
    B. Competition and Equity Markets: The Efficient Capital
        Market Hypothesis ............................................................................... 396

CONCLUSION: THE SEPARATION OF OWNERSHIP AND
AWARENESS ............................................................................................................. 401

†  Ben V. & Dorothy Willie Professor of Law, University of Iowa.

Copyright © 2009 Virginia Law & Business Review Association

373
INTRODUCTION

The idea that the large business corporation is characterized by "separation of ownership and control" will forever be identified with Adolf A. Berle and Gardiner C. Means’ *The Modern Corporation & Private Property*. The idea itself was not theirs. Already in *The Wealth of Nations*, Adam Smith wrote about business firm managers of "other people’s money" who would be unlikely to manage it with the “same anxious vigilance” shown by the active partners in a smaller firm. For the modern business corporation in the United States, the idea was well established in Thorstein Veblen’s *Absentee Ownership* in the 1920s. Nevertheless, Berle and Means’ much more approachable and graphic formulation brought the idea of separation of ownership and control into public discourse and, according to *TIME Magazine*, made their book the “economic Bible” of the Franklin D. Roosevelt Administration. *The Modern Corporation* and particularly Berle himself significantly influenced New Deal policymaking, including the passage


   Like the stewards of a rich man, they are apt to consider attention to small matters as not for their master’s honour, and very easily give themselves a dispensation from having it. Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company.

   *Id.* Smith then opined that, in foreign trade at least, joint stock companies would seldom be able to compete effectively with “private adventurers.”


of the Securities Act of 19335 and the Securities Exchange Act of 1934.6

The Modern Corporation remains a critically important historical artifact in the law of corporations and the history of the New Deal. It is widely cited to this day, although probably negatively as often as positively.7 The book is arguably the most enduring legacy of pre-war institutionalist economics. Institutionalism, with its historical and fact-intensive approach to economics, its reliance on evolutionary theories, and its distrust of markets, became the dominant economic philosophy of both the New Deal and the Legal Realists. To some degree at least, the book has suffered the same fate as Institutionalism generally—it is frequently castigated by Neoclassicists for being excessively descriptive and unscientific, and lauded by more left-leaning dissenters.

But the separation of ownership and control is hardly a distinctively institutionalist notion. It was embraced equally by marginalist Neoclassicists long prior to the publication of The Modern Corporation. The important difference was attitude. Not only for Classicists like Adam Smith, but also for Veblen, Berle and Means, and New Deal Institutionalists generally, the separation of ownership and control was a serious economic and social problem, explaining why corporations did not act in either the interests of their shareholders or the public interest.

In sharp contrast, Neoclassicists embraced the separation of ownership and control as a fundamental principle of efficient firm behavior. The principal fields of microeconomics that are external to the firm, namely price theory and industrial organization, treat the firm largely as a black box whose only goal is maximization of value. To the extent shareholder preferences differ from this goal, they almost never show up. For example, within price theory a firm sets its output or price by equating marginal cost and marginal revenue, not by consulting shareholders. Indeed, in the neoclassical model of markets, separation of ownership and control has become a virtual prerequisite to productive management and risk taking.8

While Berle and Means claimed the rhetoric of “separation of ownership

“and control,” their economic analysis suffered the same fate as institutionalist economic analysis generally. Institutionalism was overly historical and preoccupied with detail, much too descriptive, and inherently suspicious of the phenomena it was examining. Most importantly, Institutionalism was unable to devise useful theory with predictive power. The Institutionalists certainly had a point when they insisted that economics cannot be separated from social science, history, and even evolutionary biology. But in the process of attempting to incorporate everything, they gave up too much of the elegance that Neoclassicism’s much simpler forward-looking models produced. Like other institutionalist writings, *The Modern Corporation* was heavily historical and drawn to largely descriptive accounts of the dispersion of corporate ownership, as well as anecdotes about corporate power and its abuse. The empirical studies in the book pertained to the concept of separation of ownership and control. They were generally irrelevant to the conclusions that Berle and Means drew about managerial irresponsibility. By contrast, within neoclassical economics, including law and economics, separation of ownership and control has become an essential part of the analysis of the business firm and its financial structure. In addition, separation is fundamental to our assumptions about the markets in which firms operate.

The principal technical differences between the Institutionalists and the Neoclassicists lay in two things. First was the importance of Marginalism as a theory of economic choice. Second were the two schools’ profound differences concerning utility measurement and the scientific possibility of interpersonal utility comparisons.

On the first, Neoclassicism started with quantitative rules that purported to account for individual choice when resources are scarce. Under this view,

---

economic actors equate their utilities at the margin, continuously preferring things with highest incremental utility. Further, preferences were thought to be “rational” in the sense that they are transitive, and “revealed” in the sense that they are exercised through the making of observable choices, particularly in markets. This collection of observations led to a rather complete theory of prices, production, and demand and satisfaction.

By contrast, Institutionalists at least since Thorstein Veblen were deeply suspicious of Marginalism’s professed ability to explain all elements of economic behavior. Veblen found Marginalism to be both reductionist and counterfactual. It was reductionist because it purported to account for both demand and production from an entirely static picture of the economy, with no sense of movement. It was counterfactual because it did not seem to account for the complex ways that people and firms actually behave.

When basic models for the social sciences were being formed, Marginalists at the turn of the century were continuously criticized for adopting a narrow view of humanity that did not take biological evolution into account. For example, Veblen criticized marginalist economics for not being an “evolutionary” science. Marginalist economics stripped humanity down to a set of utility functions that equated human behavior with desire and completely ignored inherited characteristics. To be sure, the theory of evolution was just as reductionist, recasting desire as nothing more than the instinct to survive. But the two models developed very different mechanisms for determining appropriate social policy. The differences showed up most starkly in theories about controlling deviant behavior and criminality.

Marginalists would control crime by creating financial disincentives or limiting liberty, acting on the premise that human beings are autonomous actors who respond by degrees to various amounts of pain and pleasure. By contrast,

10. That is, if an actor prefers A over B and B over C, she must also prefer A over C.
12. Id. at 621.
13. Id. at 621–22.
Darwinians would control criminal behavior by identifying those “types” that were thought to be prone to it, and then using sterilization or other means to ensure that they could not reproduce their kind. What is often unappreciated today is the extent to which both models guided Progressive Era policymaking.  

The second important difference between the Neoclassicists and the Institutionalists lay in the value they placed on the measurement of utility and the scientific possibility of interpersonal utility comparisons. Until the early 1930s, marginalist writers generally worked from the premise that all human beings had utility functions that were more or less the same—or at least that economists could work from that assumption. As a result they could conclude that involuntary as well as voluntary transfers increased welfare. To illustrate: suppose Peter has more carrots than he needs and fewer peas. This is another way of saying that the marginal value, or utility, of the last carrot that Peter owns is less to him than the marginal value of the last pea. He would prefer to substitute carrots for peas to the point that the marginal value of the two were the same, giving him the optimal mixture. Suppose that Mary has fewer carrots than she needs and an excess of peas. She would also prefer to equate her utilities, which she could do by substituting in the opposite direction—more carrots for fewer peas. At that point, all economists, whether classical, neoclassical, or institutionalist, would agree that a voluntary exchange would be productive. By exchanging some carrots and peas with one another both Peter and Mary could arrive at points that gave them greater utility than did their starting point. That transaction would be a Pareto improvement: both Peter and Mary would be better off.

The problem occurred in situations that suggested the possibility of increased utility from an involuntary transfer. Suppose that Peter had an excess of carrots and the right amount of peas, while Mary had the right amount of peas but a shortage of carrots. Neoclassicists prior to the 1930s generally assumed that welfare, or utility, would be increased if someone simply forced Peter to give some of his carrots to Mary. After all, he had more than he


18. For information about Pareto Improvement allocations, see, e.g., Louis Kaplow and Steven Shavell, Fairness Versus Welfare 112–13 n.68 (2002).
needed, so the marginal value he placed on the excess must be quite low; she had fewer than she needed, so the marginal value she placed on them would be high. A forcible transfer would then move the carrots from a lower to a high-value use. The influential marginalist economists at Cambridge University in the late nineteenth and early twentieth century, including Alfred Marshall, Arthur Cecil Pigou, and Joan Robinson, all accepted some version of this proposition.19

Followed to its logical conclusion, this result led to socialism. The notion that people had similar utility functions that could be compared from one person to another indicated that welfare would be increased if wealth was transferred away from people who had more than they needed for basic wants, and toward people who were poor. The marginal value of a dollar to a rich person must be much less than its marginal value to someone who needs it for her next meal.

An unresolved problem, however, was whether people’s utility functions really are in fact similar. Writing in the 1930s, during the heyday of scientific Positivism and Verificationism, Lionel Robbins from the London School of Economics powerfully rejected the proposition that such comparisons were even meaningful. The Cambridge economists had simply assumed without question that people have identical utility functions, but in fact there is no empirical test one can perform in order to verify or falsify that proposition:

[S]uppose that we differed about the satisfaction derived by A from an income of £1000, and the satisfaction derived by B from an income of twice that magnitude. Asking them would provide no solution. Supposing they differed. [sic] A might urge that he had more satisfaction than B at the margin. While B might urge that, on the contrary, he had more satisfaction than A. We do not need to be slavish behaviourists to realise that here is no scientific evidence. There is no means of testing the magnitude of A’s satisfaction as compared with B’s. If we tested the state of their blood-streams, that would be a test of blood, not satisfaction. Introspection does not enable A to measure what is going on

in B’s mind, nor B to measure what is going on in A’s. There is no way of comparing the satisfactions of different people.20

With that decisive statement, “Ordinalism,” or the idea that interpersonal comparisons of utility are impossible, entered the mainstream of neoclassical economics and has for the most part been there ever since. Voluntary exchange requires no such comparison because each trader needs to know only her own utilities. For instance, Peter knows that, given what he already has, he values peas more than carrots; Mary concludes the opposite. Each makes the exchange without caring too much about how the other’s wants are satisfied. But one cannot say the same thing about the involuntary transfer: some outside observer would have to conclude that the world would be better off if Peter had fewer carrots and Mary had more.

While this discussion may seem irrelevant to the issue of corporate ownership and control, it is in fact central to understanding the way that this concept was developed within Neoclassicism. Shareholders are human beings with utility functions. Within the neoclassical framework, corporations are economic actors who have profit functions. For them, the term “utility function,” as ascribed to human preferences, is meaningless insofar as it purports to describe anything other than profits measured in a constant unit, such as dollars. In the wake of the ordinalist revolution, welfare economics and business economics, or price theory, acquired rather different focuses. Welfare economics became focused mainly on questions of general welfare, Pareto efficiency, social choice, and the derivation of a social welfare function under the constraints of Ordinalism.

By contrast, business economics largely lost its concern with utility, apart from observed individual purchasing behavior. Rather, it became preoccupied with the behavior of markets under competitive or less competitive conditions. The relevant medium of trade became currency (dollars), and dollars-as-dollars are both quantifiable and comparable as they are transferred from one actor to another. Within this model, the goal of the system is maximization of wealth and the concept of utility is largely ignored or assumed to be similar to wealth. Firms have “profit” functions, and the

Neoclassicism is a global assumption within Neoclassicism is that a firm acts rationally when it maximizes profits. Any distinctive concept of “utility maximization” of a business firm is entirely meaningless.

Further, the concept of profit maximization is indifferent to the identity and distribution of either shareholders or managers. The entire thrust of neoclassical corporate finance theory was to turn the shareholder into nothing more than an investor, who was presumed to have no interest other than the maximization of firm value, ignoring what his or her actual interest might be. The manager became nothing more than the agent of profit-maximizing decision making. Until the rise of agency cost models during the 1960s, Neoclassicism largely disregarded the ownership/control problem by positing that both the firm and its shareholders had only profit maximization in mind.21

Of course, this does not change the fact that, at bottom, all shareholders are natural biological persons,22 each with a set of desires and values that may include things other than profit maximization. Under the constraints of Ordinalism, we cannot scientifically quantify what these utility functions are, but neither can we say with certainty that anything that maximizes the value or profits of a firm necessarily maximizes the utility of its shareholders. This is hardly a problem for the neoclassical economics of the firm. It either ignores the separate utility preferences of shareholders, or assumes that the shareholders’ collective utility function is identical to the corporation’s profit function.

As this Article develops in subsequent sections, the rational behavior of a business firm—within the neoclassical model—has nothing to do with the preferences of shareholders to the extent that they prefer things that are any different from the maximization of firm value or profits. This conception of the relationship between shareholders (ownership) and managers (control) has been remarkably robust in neoclassical theory since the beginning of the twentieth century and has manifested itself in several ways, including:

- Yale economist Irving Fisher’s “separation theorem” and its offspring, which showed that the profit goals of the business

22. Some corporate shares are owned by other corporations, but ultimately a natural person, or perhaps in a few instances the government, owns the shares.
firm were completely separable from the utility functions of its diverse shareholders;\textsuperscript{23}

- Coase’s pioneering article, *The Nature of the Firm*, which showed that all of a firm’s production decisions, including those regarding its size and the extent of its integration into other markets, are entirely a consequence of a comparison of the marginal value of internal production against that of market transactions;\textsuperscript{24}

- The various corporate finance theorems of the 1950s, including the Modigliani-Miller theorem and later the efficient capital market hypothesis, which served to further disaggregate shareholders from the firm’s financial structure by treating debt and equity (ownership) as fungible, and the distribution between them as irrelevant to the firm’s value or its production decisions;\textsuperscript{25}

- Inefficiencies resulting from interest conflicts between management and shareholders were generally treated as a problem of agency costs, which are transaction or monitoring costs that are internal to the firm.

On the final point, neoclassical theory also recognized, in the guise of *agency* costs, that firms might make inefficient choices or have suboptimal financial structures. While a transaction cost is a cost of using a market, or producing an exchange between two independent actors, an agency cost is a cost of making a decision within the firm.\textsuperscript{26} While Neoclassicism saw nothing inherently inefficient in the separation of ownership and control, inefficiencies resulting from such separation would be characterized as agency

\textsuperscript{23} See *infra* notes 27–30 and accompanying text.


\textsuperscript{25} See *infra* notes 27–30 and accompanying text.

costs. One important characteristic of the great theorems of corporate finance, such as Fisher’s separation theorem or the Modigliani-Miller theorem, is that they worked in financial markets that were presumed to be perfectly competitive and without agency costs. As a result, observers could identify problems with possible legal solutions by inquiring into situations where these costs were positive. Soon after, the concept of agency costs came to perform the same role within the firm that transaction costs performed in the market—determining where legal policy could make a difference and then assigning legal entitlements in such a way so as to ensure wealth-maximizing outcomes.

I. FISHER’S SEPARATION THEOREM

Already in the first decade of the twentieth century, the brilliant Yale Neoclassicist Irving Fisher developed the early details of what was to become his “separation theorem.” Fisher’s analysis began with the notion that shareholders and firms are different. Shareholders, like all natural persons, set out to maximize their utility. They have “consumption” functions, which simply represent a list of desires, limited resources, and a set of values for prioritizing them. Firms, in contrast, set out to make profits. Indeed, a fundamental premise of both classical and neoclassical economics is that profit maximization is the goal of the firm.

If that is the case, however, then the firm’s profit function cannot be shown to be a consequence of the shareholders’ collective utility function. This is intuitively obvious, particularly when we take the ordinalist revolution into account. If shareholder utility functions are unknowable, they certainly cannot be equated with firm profit-maximization functions, which are a relatively simple matter of deciding on the course of action that produces the most profit.

27. See Douglass C. North, Comment on Stigler and Friedland: “The Literature of Economics: The Case of Berle and Means,” 26 J.L. & ECON. 269, 270 (1983) (arguing that Berle and Means addressed the problem of agency costs within the corporation long before anyone else did, and noting that Coase’s The Nature of the Firm was not published until five years later).
Fisher assumed that shareholders have utility preferences that are incapable of being specified, but nevertheless appear in individual consumption decisions. He then showed that in an efficient market for capital, a business firm will choose value maximization as a strategy regardless of shareholders’ utility preferences for dividends or reinvestment, or their preferences regarding how profits should be spent. Simply put, Fisher separation theorem states that, given perfect and complete capital markets, the production decision is governed solely by the profit-maximization objective. The production decision is separate from the consumption decision, which is governed solely by utility maximization.\footnote{For a moderately technical explanation of Irving Fisher’s theorem, see \textit{Irving Fisher’s Theory of Investment}, \url{http://cepa.newschool.edu/het/essays/capital/fisherinvest.htm}. See also \textit{Richard A. Brealey et al., Principles of Corporate Finance} 22 (McGraw-Hill Irwin 9th ed. 2007).}

Once it has accepted this formulation of the issue, the neoclassical theory of the business firm might make any of these assumptions: (1) that the firm has an aggregate “utility function” composed of the utility functions of its shareholders; (2) that the shareholders’ utility is achieved solely by the profit maximization of the firm, or that any act undertaken by the firm toward profit maximization has the unanimous consent of the shareholders; or (3) that the firm should stop caring about what the shareholders’ utility function is and consider itself a unitary profit-maximizing entity.

Neoclassical theories of the firm, of corporate finance, and even of price theory have adopted some version of assumptions (2) or (3), consistent with Fisher’s analysis. Neoclassicists who feel obliged to rationalize the complete separation of ownership and control in widely-held corporations might simply assume that shareholders want to maximize their corporation’s value. Perhaps they note that shares are held in large part by retirement funds, mutual funds, and other instruments for which maximization of value is the articulated goal. But for the most part, the formal theory of corporate finance pays very little attention to the actual preferences of shareholders insofar as they might desire the corporation to do something other than maximize its market value. In short, the theory assumed as a matter of technique that shareholders were unanimous in wanting value maximization. Until the rise of agency cost models after 1960, Neoclassicism largely disregarded the ownership/control problem by positing that both the firm and its shareholders had only profit maximization in mind.\footnote{See Demsetz, supra note 21. See also Stigler & Friedland, supra note 7, at 240.}

During the 1970s in particular, corporate finance theory was quite beset
by this assumption of shareholder unanimity that seemed inherent in the neoclassical theory of the firm. By and large, this literature did not sample actual shareholder preferences to see if they were truly unanimous. Rather, it simply assumed that all shareholders were profit maximizers, although they might have had different attitudes about risk. But for the most part, in corporate finance theory, the distinctive preferences of shareholders are simply ignored. That which maximizes the value of the firm is all but conclusively presumed to be what the shareholders prefer, and that is the end of the matter. Indeed, actual polling of shareholder preferences on most production and finance decisions becomes a costly, frustrating waste of time and resources. As Eugene Fama and Merton Miller said in the 1970s:

Where there exist organized capital markets in which shares can be freely bought and sold and where these markets are perfect . . . it is possible to develop an objective, operational decision criterion for management that (1) does not involve stockholder utility functions directly but (2) leads to precisely the same investment and operating decisions that each stockholder would make if he were running the firm himself.  

The basic logic of Fisher’s separation theorem was that the goal of the firm is always to maximize overall returns, thus giving the shareholders the ideal opportunity to spend the profits as they please. Today, Fisher’s


33. Eugene F. Fama & Merton H. Miller, The Theory of Finance 69 (1972). For an enthusiastic embrace of separation of ownership and control as permitting specialized decision making and reducing agency costs, see Fama & Jensen, supra note 8, at 301.

separation theorem is regarded as a building block for the more general Modigliani-Miller theorem of corporate finance. This theorem, developed in the 1950s, states that in an efficient market for capital, a firm’s value is not a function of the way it is financed (i.e., its ratio of debt to equity). In this model, the number, identity, or interests of shareholders become entirely irrelevant when the conditions of the theorem are satisfied.\textsuperscript{35} Of course, this model, like all neoclassical models, is severely reductionist. It assumes that profit is the only thing that shareholders want. To the extent an individual shareholder believes the firm should be pursuing some goal inconsistent with profit maximization, his or her wishes are simply ignored.

Assumptions such as Fisher’s separation theorem became particularly important in the analysis of highly competitive markets, where profit-maximizing behavior is an essential prerequisite to firm profitability. That is, the more competitive the market in which a firm operates, the more important it becomes to ignore the divergent wishes of shareholders.\textsuperscript{36}

Ironically, this view stated the Berle and Means concern somewhat backwards. For them, separation of ownership and control was a problem of bigness and, at least to a degree, of monopoly. But in fact, substantial market power creates more tolerance for firm discretion than does competition. In other words, for the widely-held firm in a highly competitive market, separation of ownership and control becomes a matter of survival.

II. VALUE MAXIMIZATION AND THE NATURE OF THE FIRM

Following Fisher, the next great statement of the neoclassical business firm was Ronald Coase’s \textit{The Nature of the Firm}, which was published in 1937 while Coase was still at the London School of Economics.\textsuperscript{37} Coase simply assumed what Fisher had labored to prove. While Coase’s article discussed decision making in several large, widely-held American firms, it never so much as mentioned the diverse desires of shareholders. Coase began with the premise that the firm invariably seeks to maximize its profit, and queries how its managers will decide the size of the firm and the number of markets in

\textsuperscript{35} See supra notes 29–30 and accompanying text. See also Demsetz, supra note 21, at 375 (explaining that corporate performance does not depend on identity or configuration of shareholders); Harold Demsetz & Belen Villalonga, Ownership Structure and Corporate Performance, 7 J. CORP. FIN. 209, 210 (2001) (stating that shareholders will eventually adopt the ownership form that maximizes returns).

\textsuperscript{36} On this point, see Fama & Jensen, supra note 8.

\textsuperscript{37} Coase, supra note 24, at 404.
which it will operate. Perhaps the best thesis statement of The Nature of the Firm is the one that Coase himself gave a half century later:

Whether a transaction would be organized within the firm . . . [or integrated] or whether it would be carried out on the market by independent contractors depended on a comparison of the costs of carrying out these market transactions with the costs of carrying out these transactions within an organization, the firm.  

In making any production or distribution decision, from whether to switch suppliers, or not to use a supplier at all, a firm relentlessly and continuously compares costs and payoffs, always taking the course that is most profitable. The Coasian firm never consults shareholders except insofar as one of them might have some information or expertise to offer. If the shoe manufacturer considers whether to continue purchasing laces from an independent supplier or to integrate vertically into lace production by building its own lace factory, it simply compares the costs and benefits of each alternative. Importantly, these decisions are not simply “operational,” they are also “structural.” The aggregation of all of these decisions establishes how large the firm is “horizontally,” how many different products it produces, and also how large it is “vertically,” in the sense that it integrates different steps in the manufacturing and distribution process.

Coase has repeatedly professed that The Nature of the Firm was based on hard empirical work in the field. To be sure, he does not claim that his article was based on quantitative methods. Rather, Coase claims that he formulated his ideas by examining the details of how production decisions are made in real life firms. Throughout his career Coase wrote with irritation about armchair economists who think of economics in terms of “markets” or

40. See e.g., Coase, supra note 38.
41. To the contrary, in speaking of his enlightening visit to a Ford Motor plant he observed that people were more willing to talk to him because “I do not want statistics. All I want are statements that are suggestive from the point of view of fitting into a theory of integration.” Coase, supra note 38, at 14.
42. Coase, supra note 38.
“functions,” concepts that are abstracted from reality. Much of Coase’s work, particularly The Problem of Social Cost, has been made the subject of empirical studies. Further, in his 1991 Nobel Prize lecture he railed at economists whose principal work consisted of a set of abstractions.


I don’t reject any policy without considering what its results are. If someone says there’s going to be regulation, I don’t say that regulation will be bad . . . . What we discover is that most regulation does produce, or has produced in recent times, a worse result. But I wouldn’t like to say that all regulation would have this effect because one can think of circumstances in which it doesn’t . . . . Almost all the studies—perhaps all the studies—suggested that the results of regulation had been bad, that the prices were higher, that the product was worse adapted to the needs of consumers, than it otherwise would have been. I was not willing to accept the view that all regulation was bound to produce these results. Therefore, what was my explanation for the results we had? I argued that the most probable explanation was that the government now operates on such a massive scale that it had reached the stage of what economists call negative marginal returns. Anything additional it does, it messes up.


This neglect of other aspects of the system has been made easier by another feature of modern economic theory—the growing abstraction of the analysis, which does not seem to call for a detailed knowledge of the actual economic system or, at any rate, has managed to proceed without it. Holmstrom and Tirole, writing on The Theory of the Firm in the recently published Handbook of Industrial Organization, conclude at the end of their article of 63 pages that “the evidence/theory ratio . . . is currently very low in this field.” Peltzman has written a scathing review of the Handbook in which he points out how much of the discussion in it is theory without
Notwithstanding these protests, *The Nature of the Firm* is an exercise in pure theory. As Coase wrote:

It is hoped to show in the following paper that a definition of a firm may be obtained which is not only realistic in that it corresponds to what is meant by a firm in the real world, but is tractable by two of the most powerful instruments of economic analysis developed by Marshall, the idea of the margin and that of substitution, together giving the idea of substitution at the margin.46

Why does a firm producing an undifferentiated product in competition commit its resources to producing a second product, instead of continuously producing more and more of the original product? Coase’s answer was that “there may be a point where it is less costly to organise the exchange transactions of a new product than to organise further exchange transactions of the old product.”47 Given that both markets and internal production have costs, a firm maximizes its profits by choosing among the most profitable

47. *Id.* at 402. In *The Nature of the Firm*, Coase used “marketing costs” to describe what would later be characterized as transaction costs: “To determine the size of the firm, we have to consider the marketing costs (that is, the costs of using the price mechanism) . . . .” *Id.* at 403. See also *Id.* at 405 (describing how the businessman combines the functions of “initiative,” which refers to use of the market, and “management,” which refers to supervision of internal production, and that the combination is the result of “marketing costs”). Fifty years later, when recalling the writing of *The Nature of the Firm*, Coase spoke of the same costs as transaction costs: “The solution [to the problem of understanding why firms integrate vertically rather than purchasing] was to realize that there were costs of making transactions in a market economy and that it was necessary to incorporate them into the analysis.” *See* Coase, *supra* note 38.
alternatives, continuously comparing the marginal costs and returns of doing something “in-house” against the costs and returns of purchasing from or selling to someone else.

To be sure, Coase made several observations to the effect that purchasers in companies are continuously searching for the best deal. 48 He also expanded this to include the idea that a firm might decide to manufacture one’s own shoe laces when the market conditions for purchasing them are too unfavorable. But this is largely common sense and common knowledge, and certainly does not establish the elegant theory of production choice that Coase developed in *The Nature of the Firm*. The theory underlying Coase’s brilliant paper is no more empirical than the observations of the early Marginalists such as Jevons and Marshall, that individuals have declining marginal utility and equate their utilities at the margin. 49 One might observe that the home shopper who has three dozen eggs in the refrigerator but no milk will go to the store to purchase milk and not eggs. But that observation tells us little about the theory that consumers equate their utilities at the margin.

Toward the end of his article Coase queried to himself and his readers whether his concept of the firm “fits in with that existing in the real world.” 50 He did so, not by looking at firms, but rather by examining the legal relationship of master and servant. After quoting a lengthy definition from an English jurist, Francis Raleigh Batt, 51 Coase concluded that what the employee (servant) and agent (buyer or reseller) had in common was not the presence of a fixed wage or commission, but rather the degree of freedom that an agent could exercise. 52 Coase concluded:

When we are considering how large a firm will be the principle of marginalism works smoothly. The question

---

48. See Coase, supra note 38, at 8–10, 13 (recounting a visit to the United States prior to writing *The Nature of the Firm*, where he visited the purchasing department at a Union Carbide plant and listened to telephone conversations that “gave [him] a lively sense of the possibilities of substitution”; also recounting discussions at General Motors about the acquisition of Fisher Body works, its former trading partner, and being told that it was to ensure that Fisher’s plant would remain located close to GM).


50. Coase, supra note 24, at 403.


52. Id. at 404 (quoting Batt, supra note 51, at 7).
always is, will it pay to bring an extra exchange transaction under the organising authority? At the margin, the costs of organising within the firm will be equal either to the costs of organising in another firm or to the costs involved in leaving the transaction to be “organised” by the price mechanism. Business men will be constantly experimenting, controlling more or less, and in this way, equilibrium will be maintained.53

The Nature of the Firm did for firm structure what Fisher’s separation theorem had suggested for firm finance: it turned the firm into an engine whose only goal was the pursuit of profits, which it achieved by continuously comparing the marginal costs and benefits of doing things in different ways. The result was a “moving equilibrium,” within which managers made continuous decisions about how and what to produce, and how and what to purchase. At the margin, internal and external costs are equalized. Significantly, however, this was a set of purely technical problems, resolved for the manager by experience and the economist, by price theory and industrial organization. To the extent they might differ, the independent wishes of shareholders had no place.

III. THE FINANCIAL STRUCTURE AND VALUE OF THE CORPORATION

The earliest neoclassical theorems in corporate finance, such as Irving Fisher’s separation theorem, assumed that capital markets were efficient.54 In fact, the roots of the modern efficient capital market hypothesis were developed in neoclassical marginalism early on. As the early Marginalists observed in the late nineteenth century, people tend to equate their utilities.55 They purchase a good until the marginal utility of that good declines to the level they experience for some other good.56 The corollary in finance is that people equate their returns. Stocks become investment vehicles whose prices are calculated to produce the same level of return, once adjusted for risk.

53. Coase, supra note 24, at 404.
54. See supra notes 29–30 and accompanying text.
One important difference between classical and neoclassical value theory lay in the treatment of risk and uncertainty. Because Classicists measured value by looking at past averages, the theory did not explicitly incorporate the risk of uncertain future events. Things such as the value of labor or the value of a business firm were measured by reference to previous investment; risk of future events did not formally fit into the theory. To be sure, business persons investing in the nineteenth century certainly took anticipated risks into account, but the classical value model did not account for them. In very sharp contrast, Marginalism’s criteria of willingness-to-pay, or expected value, almost always involved a certain amount of uncertainty. For longer-run investments or less stable markets the uncertainty could be considerable. Determining how to accommodate uncertainty about the future into economic modeling proved to be a central problem of neoclassical economics in the first half of the twentieth century.

Before modern corporate finance theory could emerge, several things had to be worked out. First, Marginalism had to develop a robust theory of competition because of its forward-looking nature, and because it was a theory in which information, risk, and uncertainty acquired heightened importance. Second, this theory had to be applied to the corporate equity market. The theory, coupled with a set of empirical studies of commodity and stock market behavior, led to the formulation of the efficient capital market hypothesis.

Marginalist economics initially created a great deal of doubt about the competitiveness and even the robustness of markets. Also, many of the earliest Marginalists abandoned the commitment to free markets that was explicit in classical political economy. Some even toyed with socialism as an alternative to free markets. Prominent Neoclassicists backtracked considerably from the classical hostility, toward economic regulation. Major


59. See infra notes 70–88 and accompanying text.


61. See WILLIAM S. JEVONS, THE STATE IN RELATION TO LABOUR 1–21 (MacMillan & Co. 1882). See also HENRY SIDGWICK, THE PRINCIPLES OF POLITICAL ECONOMY 518–44 (MacMillan & Co. 1887) (criticizing the laissez-faire position of the Classicists). On the
technical controversies within neoclassical economics served to create significant doubts about the efficiency of markets.\(^2\)

Gradually, Neoclassicism was able to formulate the details of a more or less robust model of competition, even though the domain of so-called “perfect competition” within Marginalism was never as broad as the strenuous competition that Classicists believed prevailed in nearly every market. The Neoclassicists had to deal with numerous complexities that the marginalist model contemplated, such as increasing returns to scale, which gave larger firms a cost advantage over smaller ones.\(^3\) Marginalism also led to theories of imperfect competition and product differentiation, both of which made marginal cost pricing unworkable.\(^4\) In addition, marginalist corporate finance theory had to work out some important problems regarding the relationship between a business firm’s market incentives, its selection of capital sources, and the potentially separate incentives of its stockholders. The Fisher separation theorem discussed above was an important first step; the great corporate finance theorems of the mid-twentieth century were another.

A. Modigliani-Miller: The Fungibility of Ownership and Debt

Today, Irving Fisher’s separation theorem is regarded as a building block for the more general Modigliani-Miller theorem of corporate finance. Developed in the 1950s, the theorem states that in an efficient market for capital, a firm’s value is not a function of the way it is financed (i.e., its ratio of debt to equity). In this model, the number, identity, and interests of shareholders become irrelevant to the firm’s financial decisions when the conditions of the theorem are satisfied.\(^5\)

In the traditional law and theory of business firms, ownership and debt


\(^{63}\) See Henry C. Adams, Relation of the State to Industrial Action, 1 PUBLICATIONS OF THE AMERICAN ECONOMIC ASSOCIATION 7, 52, 59–64 (Am. Econ. Ass’n 1887) (arguing that industries subject to significant scale economies must be regulated by the government); Hovenkamp, supra note 58, at 15–16. See also Hovenkamp, supra note 62, at ch. 23.

\(^{64}\) See CHAMBERLIN, supra note 62, at ch. 5.

were regarded as two very distinctive things. If the firm was a corporation, ownership was a set of property rights in the form of stock, and owners made the operational decisions. By contrast, debt was contractual and gave creditors a right to collect periodic interest payments and eventually reclaim their principal. But ordinarily, debt gave them no say in the firm’s operations. A business firm commonly has both owners and creditors and also has considerable control over its debt-to-equity ratio. For example, if a firm is thinking of building a new plant, it might finance the project by issuing additional shares (ownership) or by borrowing money.

Under traditional notions of corporate finance, a firm could reduce its cost of capital and thus increase its value by taking on debt. Typically, debt is cheaper because the rate of return on bonds is lower than the returns to capital, reflecting the lower risk of bonds. However, if a firm borrows too much—that is, if it becomes excessively leveraged—then risk begins to rise, and both stockholders and bondholders will demand higher returns, causing the cost of capital to rise again. As a result, under traditional corporate finance theory, the cost of capital falls as the corporation takes on more debt, bottoms out at a certain point, but rises again as the amount of debt is regarded as excessive. One important way of maximizing the firm’s value by reducing its cost of capital is to find the optimal debt-to-equity ratio. Traditionally, both too little and too much debt were regarded as costly.

The Modigliani-Miller theorem was named after its developers. Franco Modigliani was an Italian-born Jewish economist who fled Europe during the Italian Fascist purge. He eventually settled in the United States and made his career primarily at the Massachusetts Institute of Technology. Merton Miller was an American economist who spent most of his senior career at the University of Chicago. The Modigliani-Miller theorem was worked out in a series of articles written in the late 1950s and early 1960s. The theorem states that in a perfectly functioning market for capital, debt and equity are completely fungible. As a result, it does not matter what proportion of debt to equity a firm has; its value is determined entirely by its stream of earnings in the markets where it operates.

---


67. See FRANCO MODIGLIANI, 3 THE COLLECTED PAPERS OF FRANCO MODIGLIANI, at xiii (A.
Modigliani and Miller reasoned that in a perfectly functioning capital market, all participants would be able to optimize. For example, suppose there are two firms, Xequity and Xmix, which are identical in every way except that Xequity’s entire capital comes from stock ownership, while Xmix’s capital is based 60% on equity and 40% on debt. How much would a prospective owner be willing to pay for each of these firms? The traditional answer was that Xmix would be worth either more or less than Xequity depending on how close it was to the optimal debt-to-equity ratio.

By contrast, the Modigliani-Miller analysis showed that, where their assumptions were met, the value of the two firms would be precisely the same. This is because, in a perfectly functioning capital market, participants could make their own offsetting debt-to-equity substitutions. For example, an investor could purchase the debt-free firm but borrow 40% of the money, leaving herself in a debt-equity position precisely equal to that of the leveraged firm’s position. As a result, if one firm was priced more highly than the other, the investor would take the cheaper one without regard to the debt-to-equity ratio. Any time the market tended to value one of the two firms higher than the other based simply on its debt-to-equity ratio, rational investors would borrow a complementary amount of money, purchase the cheaper option, and trade the value back to the equilibrium point. In equilibrium the investor would face the same risk for any mix of stock and equity. As a result, equity-to-debt would have no impact on the value of the firm. Under this concept, stock ownership, or equity, becomes nothing more than one of two completely interchangeable ways of providing capital to the firm.

The theorem applies only to perfectly competitive financial markets without transaction costs. As a result, it comes with the same caveats as the Coase theorem. That is, although theorizing a world of perfect, costless

Abel ed., 1980), stating the following as the basic theorem:

[W]ith well-functioning markets (and neutral taxes) and rational investors, who can “undo” the corporate financial structure by holding positive or negative amounts of debt, the market value of the firm—debt plus equity—depends only on the income stream generated by its assets. It follows, in particular, that the value of a firm should not be affected by the share of debt in its financial structure or by what will be done with the returns—paid out as dividends or reinvested (profitably).

See also Demsetz, infra note 21, at 386 (arguing that corporate performance does not depend on identity or configuration of shareholders); Demsetz & Villalonga, supra note 35, at 211–12 (demonstrating that shareholders will eventually adopt the ownership for maximizing returns); Michael Klausner, The Contractarian Theory of Corporate Law: A Generation Later, 31 J. CORP. L. 779 (2006).
capital markets is informative, corporate finance policy choices must be considered in the context of the real world, where transaction costs are never free. For example, to the extent that debt receives tax treatment that is more favorable than the treatment given to dividends, the impact will be to shift firms more toward debt. This could result in excessive leverage and fixed interest cost commitments that cannot be paid in the event of a financial downturn.

The Modigliani-Miller theorem can also be seen as a specific application of Coase’s idea in The Nature of the Firm—that a firm’s managers will always choose the more profitable course of operation. Modigliani-Miller extends this observation to choices about the firm’s financial structure. In the absence of transaction or agency costs, the value of the firm is invariant to the way that it is financed—either by increasing ownership through equity or by increasing borrowing through debt. The corollary is that in the presence of such costs, the firm will select the more profitable alternative. As a result, the firm’s choice about whether to issue more shares or take on more debt is driven entirely by the presence of frictions in the system, such as taxes, agency problems, other internal inefficiencies, or imperfect information, but not by any notion that the firm is inherently more valuable under one or another form of organization. Further, its choice of a debt-to-equity ratio is entirely independent of the individual wishes of shareholders.

B. Competition and Equity Markets: The Efficient Capital Market Hypothesis

As the modern neoclassical model of perfect competition developed through the first half of the twentieth century, the role of information became increasingly important. Perfect competition depended on markets with a fairly large number of buyers and sellers, a lack of significant scale economies, and the free flow of information. In 1921, University of Chicago economist Frank Knight identified costless flow of information as a precondition to effective competition in his important book Risk, Uncertainty and Profit. In addition, Knight introduced the concepts of risk and certainty as inherent in Marginalism’s emphasis on reasonable expectations. For Knight, “risk”

68. See supra notes 45–48 and accompanying text.
70. KNIGHT, supra note 56, at 78–87.
referred to future variations whose probabilities were knowable. With accurate foreknowledge of probabilities, risks could be traded under competitive conditions. For example, a ten percent chance of making a $1000 oil discovery is worth $100. In contrast, “uncertainty” referred to future events whose probabilities could not be known. In such cases, investors would demand a premium as compensation for exposure to an adverse unpredictable outcome.71

Many of the early Marginalists viewed the stock market with suspicion, regarding it as not conforming to the usual laws of supply and demand. Certainly, the boom-bust stock price cycles of the late nineteenth and early twentieth century frustrated the application of basic competition theory to stock pricing. Prices appeared to fluctuate wildly, with no apparent relation to the value of the underlying firm.72 Under this line of thinking, technical analysis flourished, with stock traders hoping to pick winners by searching for predictable similarities in past pricing behavior.73

But a more theoretical purist strand in neoclassical economics developed the view that, notwithstanding the frenzy with which stocks are often purchased and sold, overall pricing tends to reflect fundamental values. For example, Irving Fisher, author of the separation theorem,74 consistently argued that stock prices reflected intrinsic values, in which returns to stocks operated as an “implied” rate of interest and owners were compensated with higher returns in exchange for taking on greater risk.75

---

71. Id.
72. See, e.g., John Maynard Keynes, General Theory of Employment, Interest and Money 156 (1936) (arguing that the stock market operates as a kind of “beauty contest” in which shares’ prices were based not on fundamental value, but on each buyer’s prediction of what valuation others would place on a firm’s shares). See also John Hicks, Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory (1939); Robert Rhea, The Dow Theory (1932).
74. See supra notes 29–30 and accompanying text.
75. Irving Fisher, The Rate of Interest: Its Nature, Determination and Relation to Economic Phenomena 10 (1907) (speaking of an “implied rate of interest” in stocks that reflects the investor’s anticipation of returns). See also id. at 216 (noting that the intrinsic value of stocks is such as to produce a higher rate or return because stocks are also accompanied with more risk). More than two decades later, Fisher returned to the same themes in Irving Fisher, The Theory of Interest: As Determined by the Impatience to Spend Income and Opportunity to Invest It 69, 126–27 (1930). See also Benjamin Graham & David L. Dodd, Security Analysis: Principles and Technique 15–18, 404–11 (1951) (emphasizing a strategy of “value investing” by studying fundamentals in search of undervalued stocks).
also insisted that the price of shares reflected the intrinsic value that they represented—namely, objectively reasonable expectations of future earnings and dividends.76 Mathematically, the value of a corporation is the expected value of its stream of future earnings. Building on Williams’ work, University of California economist Harry Markowitz then developed the idea that an optimal portfolio of stocks consists of stocks of differing risk levels, and riskier investments are offset by higher rates of return, notwithstanding higher variability.77

The efficient capital market hypothesis was constructed on the marginalist theory of perfect competition, where every market participant is a price taker, and the price of a stock quickly moves toward an equilibrium that tends to equalize its risk-adjusted return to that of other stocks and financial instruments.78 The theory developed in stages. First was the observation that returns at the margin will be equalized. Next was the observation that, to the extent the market discounts all information about a stock into its price, the current price is always the “correct” one. Finally, it was observed that even high-risk and low-risk stocks should produce the same return in the long run because high-risk stocks will be compensated through a stock price that yields a higher return.79 As a result, any randomly selected mixture of stocks should perform just as well as any other similarly diversified mixture.

From that point the only missing ingredient was informational efficiency, or the idea that the market price of a security is a reflection of the information that is publicly known about it. To the extent that information is both accurate and relatively quickly disseminated, the market price will tend to reflect rational expectations about fundamental value.80 Already in 1900

76. JOHN BURR WILLIAMS, THEORY OF INVESTMENT VALUE 54–58 (1938).
77. See Harry Markowitz, Portfolio Selection, 7 J. FIN. 77 (1952).
78. See Jean-Jacques Laffont & Eric S. Maskin, The Efficient Market Hypothesis and Insider Trading on the Stock Market, 98 J. POL. ECON. 70, 70–71 (1990) (noting that the efficient capital market hypothesis assumes nearly perfect competition and breaks down under oligopoly, where prices and the release of information may be strategic—if transaction costs are positive or there are serious asymmetries in information then various versions of the hypothesis may not apply).
79. Markowitz, supra note 77.
Louis Bachelier, a French mathematician, had written a doctoral dissertation entitled *The Theory of Speculation*, arguing that the history of commodity prices shows that they are in fact randomly distributed, making it impossible to predict future prices from past price histories.\(^8^1\) Beginning in the 1930s, a number of studies suggested the same thing for stock prices.\(^8^2\) For example, detailed recording of a series of throws of a single die provides information that there is a one-in-six chance of getting a five, but no sequence in historical throws provides any useful information about predicting a sequence in future throws. As a result, an efficient investor can forget the historical research about price movements.

In his now famous doctoral dissertation, Eugene Fama assembled this theory about competition and information dispersion, as well as the empirical studies of pricing behavior into what is now known as the efficient capital market hypothesis. This hypothesis generally states that in any market in which information flows without restraint, current market prices reflect investors’ collective beliefs about the value of the goods that are being traded. While the efficient capital market hypothesis can be applied to any market that satisfies its conditions, and has frequently been applied to commodities markets,\(^8^3\) its main impact has been in the analysis of stock market pricing.

The efficient capital market hypothesis comes in three versions: weak, semi-strong, and strong.\(^8^4\) The weak form states that current prices reflect all of the information contained from observations of previous investment prices. As a result, historical pricing information is not useful for predicting future prices; thus so-called “technical” analysis from price movements is useless as a predictor of future prices. Under the semi-strong form, current prices reflect all public information, including not only technical information

---


\(^8^4\) See Fama, supra note 80, at 388 and accompanying text.
but also information pertaining to “fundamentals,” which includes all information about the performance and prospects of a firm. These “fundamentals” include a firm’s assets and liabilities, price-to-equity ratio, and other important financial ratios. Information discussed in the newspaper has already been reflected in the market price, and trading on that information is of no use. Further, studying a firm’s fundamentals in order to identify under- or over-valued stocks is pointless; thus neither technical analysis nor fundamental analysis will help predict a stock’s future price. Finally, the strong form enhances the previous form in that even private information is reflected in the stock price. As a result, even information from insider trading will be included.85

Both the strong and the semi-strong versions of the efficient capital market hypothesis have strong policy implications for corporate disclosure and finance. Principally, they mitigate strongly against hard regulation but in favor of disclosure of information.86 With respect to information, mandatory disclosure is more important for smaller companies than for larger publicly traded companies that are likely to be followed by a large number of analysts.87 In general, the amount of regulation of information that should be supplied varies inversely with the amount of information available and disseminated by private analysts.88 Finally, the type of financing a firm chooses and its production or expansion decisions will always be reflected in the market price, making command-and-control regulation largely unnecessary.

CONCLUSION: THE SEPARATION OF OWNERSHIP AND AWARENESS

Nearly every interesting assertion in Berle and Means’ Modern Corporation

88. Id. at 812.
has been disputed, including separation of ownership and control. Stigler and Friedland disputed the hypothesis that ownership and control were substantially separated in any functional sense. Stigler and Friedland also disputed the notion that ownership and control were substantially separated in any functional sense. Other scholars have suggested that Berle and Means' suggestion that stock ownership prior to the 1930s was as dilute as they believed. These scholars also believed that even in widely-held corporations, organized blocks of stockholders preserve effective ownership control. Other literature disagrees strongly and gives evidence that, at least as a factual matter, Berle and Means were correct. In many large corporations stock ownership is in fact diffuse, and managers are not effectively controlled by owners. Regardless, the separation of ownership and control that Berle and Means claimed to identify would be a crucial component in the formation of corporate governance and finance theory for the balance of the twentieth century.

Of course, in an important sense, the historical fact of separation of ownership and control is but a detail. Much more important are the conclusions to be drawn. For Berle and Means and other Institutionalists, Legal Realists, and more recent critics such as Ralph Nader, the result was corporate autocracy and waste. For the Neoclassicists, by contrast, the result was an essential prerequisite to theory. By the 1970s, Marginalism and price

———

89. Stigler & Friedland, supra note 7, at 244.
91. Clifford G. Holderness, The Myth of Diffuse Ownership in the United States, 22 REV. FIN. STUD. 1377 (2009) (noting that 96% of a sample of publicly traded firms in the United States have blockholders that, on average, own 39% of the shares in their corporation).
theory came to define every aspect of the business firm, including its behavior, its market structure, and its financial structure. Stock ownership became no more than a detail and, even when it was relevant, did no more than explain deviations from the norm.

Therefore, the result of marginalist finance theory, particularly the efficient capital market hypothesis, was to take Berle and Means’ separation of corporate ownership and control one step further, to the separation of ownership and awareness. In an efficient capital market, investors can maximize their returns without knowing the products that a firm produces, the markets in which it operates, the makeup of its board of directors, or even the name of its CEO. A random selection of stocks produces the same return as the most careful research.95 Indeed, under the strong version of the efficient capital market hypothesis, the actions of managers become irrelevant because these actions will immediately be reflected in the stock price. The effect of these theories on the shareholder was significant. Shareholders of publicly traded corporations could invest with indifference and indiscrimination—a massive shift away from the nineteenth century vision of the corporation as a device to facilitate investment predominately by groups of active owner-operators who sought to limit their liability.