Coase, Institutionalism, and the Origins of Law and Economics†

HERBERT HOVENKAMP*

INTRODUCTION ...................................................................................................... 499
I. BRITISH MARGINALISM BEFORE COASE ............................................................. 502
   A. MARGINALISM, EQUILIBRIUM, AND THE COST OF MOVING RESOURCES .. 502
   B. THE ORDINALIST REVOLUTION ................................................................. 508
II. THE RISE OF INSTITUTIONALISM........................................................................ 515
   A. DARWIN, BEHAVIORISM, AND THE REVOLT AGAINST MARGINALISM ..... 515
   B. INSTITUTIONALIST ECONOMICS FROM VEBLEN TO THE LEGAL REALISTS ...... 521
III. COASE: FROM NEOCLASSICISM TO NEW INSTITUTIONALISM ........................... 529
   A. COASE AND OLD INSTITUTIONALISM ....................................................... 529
   B. THE BUSINESS FIRM AND INSTITUTIONALISM, OLD AND NEW ................. 530
   C. LAWRENCE KELSO FRANK ....................................................................... 532
   D. CLARK AND OVERHEAD COSTS ............................................................... 532
   E. PIGOU, KALDOR, AND ROBINSON: THE FIRM IN EQUILIBRIUM ................. 537
CONCLUSION: FROM COASEAN INSTITUTIONALISM TO LAW AND ECONOMICS ...... 540

INTRODUCTION

Ronald Coase transformed our understanding of the role of transaction costs in the economic and legal systems. In a way, it can be said that he invented the modern discipline of law and economics.1 He was also a creature of his times. His education was steeped in British and Continental economics and law.2 He studied at the London School of Economics (LSE) in the late 1920s and early 1930s, during its heyday under the leadership of Lionel Robbins.3 The London School of Economics had been founded in 1895, making it a very young upstart among British universities. The dominant British economic thinking of the time was coming mainly from Cambridge University. In large part, Robbins’s leadership of LSE in the 1930s was dedicated to breaking the hold that the Cambridge School

† Copyright © 2011 Herbert Hovenkamp.
* Ben V. & Dorothy Willie Professor, University of Iowa College of Law.
3. Lord Lionel Robbins was chair of political economy at LSE from 1929 until 1961. Coase described Robbins as “the most influential figure of all” at LSE in the 1930s. COASE, ECONOMICS AT LSE, supra note 2, at 211; see also WILLIAM H. BEVERIDGE, THE LONDON SCHOOL OF ECONOMICS AND ITS PROBLEMS, 1919–1937 (1960); ELIZABETH DURBIN, NEW JERUSALEM: THE LABOUR PARTY AND THE ECONOMICS OF DEMOCRATIC SOCIALISM 101–36 (1985); LIONEL ROBBINS, AUTOBIOGRAPHY OF AN ECONOMIST 105–31 (1971).
(chiefly Alfred Marshall, Arthur Cecil Pigou, and later, Joan Robinson) had over neoclassical economic theory in the early twentieth century.4

Except for two periods away, Coase remained at LSE until he emigrated to the United States in 1951.5 During the time Coase was at LSE, the school was forging a name for itself in the development of economic theory and welfare economics, under such scholars as Robbins, John Hicks, Paul Sweezy, Abba Lerner, Nicholas Kaldor, and Tibor Scitovsky.6 By his own admission, Coase spent more time studying law than economics, and his obsession was figuring out how institutions in the real world operate.7

Coase has occasionally identified himself with the group of economists called “institutionalists,”8 although there is scant evidence that he took any of the first generation of institutionalists very seriously. Today, however, many of Coase’s followers describe their discipline as “new institutional economics.”9


5. Coase took a Bachelor of Commerce degree at LSE in 1932 and was on its teaching staff from 1935 to 1951, except for an assignment as a government statistician during World War II; he also made a visit to the United States in 1931 and 1932, where he collected observations for his article The Nature of the Firm. See R.H. Coase, The Nature of the Firm: Origin, 4 J.L. Econ. & Org. 3 (1988) [hereinafter Coase, Nature of the Firm: Origin].

6. For a good history of the principal faculty and most famous students at LSE, see the website at http://homepage.newschool.edu/het/schools/lse.htm. For John Hicks’ perspective, see John Hicks, Introductory: LSE and the Robbins Circle, in 2 Money, Interest and Wages: Collected Essays on Economic Theory 3 (Harvard Univ. Press 1982).

7. See Coase, Nature of the Firm: Origin, supra note 5, at 6 (recalling that in his years at LSE as a student he took no courses in economics and spent most of his time studying law, particularly “industrial law”).


Although the diversity of its adherents makes a definition elusive, “institutionalism” historically refers to a group of mainly American economists whose work stretched from the beginning of the twentieth century until after the New Deal.\textsuperscript{10} Their place in economic theory is somewhat outside the mainstream,\textsuperscript{11} but they have recently experienced a resurgence with the rise of interest in behavioral economics and socioeconomics.\textsuperscript{12} This first generation of institutionalists emphasized the importance of human-created institutions that serve to allocate power or resources, the rules that these institutions develop and employ, and their effect in the overall economy. The old institutionalists generally rejected the neoclassical notion that preferences (value) must simply be accepted as asserted. Most of them were far more interested in examining the sources of human preferences, emphasizing their links with either evolutionary biology or behaviorist psychology. Unlike mainstream neoclassicists, who tended to reduce economics to the study of markets, the institutionalists believed that voluntary market exchange is only one of many institutions that move resources through society.\textsuperscript{13}

This paper examines the relationship of Ronald Coase’s thought to neoclassicism and institutionalism. The first generation of institutionalists rejected or severely qualified marginalist analysis, as well as the emergent neoclassical creed that the study of naked individual preference is the exclusive methodology of economic science. They came to believe that in a world in which resources are scarce and their movement is costly, a variety of institutions emerge for determining the course of movement by forcing individuals to make tradeoffs.\textsuperscript{14}

Coase’s most important work seemed to merge neoclassicism with certain elements of institutionalism by incorporating marginalist analysis into the study of institutions.\textsuperscript{15} Like other neoclassicists, he was not concerned about the source of preferences but only with the mechanisms by which they are asserted.\textsuperscript{16} More explicitly, by recognizing individual preference orderings and market exchange as the principal movers of resources, Coase reduced the problem of resource movement to one of “transaction costs.”\textsuperscript{17} While he barely acknowledged the


\textsuperscript{11} Nonetheless, the old institutionalists continue to publish journals, including \textit{The Journal of Economic Issues} and the \textit{Journal of Institutional Economics}. Several scholarly organizations support their work, including the Association for Evolutionary Economics. \textbf{See AsS’N FOR EVOLUTIONARY ECON.}, http://www.associationforevolutionaryeconomics.org/division.php?page=institutional_economics&sub=associations.

\textsuperscript{12} \textbf{See}, e.g., \textbf{The Journal of Socio-Economics}, http://ideas.repec.org/s/cee/soceco.html.

\textsuperscript{13} \textit{See infra} text accompanying notes 123–24.

\textsuperscript{14} \textit{See infra} Part II.

\textsuperscript{15} \textit{See infra} text accompanying notes 162–69.

\textsuperscript{16} This statement seems quite true, notwithstanding that in an essay on new institutional economics, Coase specifically acknowledged “the work of our colleagues in law, anthropology, sociology, political science, sociobiology, and other disciplines.” Coase, \textit{New Institutional Economics, supra} note 8, at 72.

\textsuperscript{17} \textit{See infra} text accompanying notes 168–69.
existence of the first generation of institutionalists, Coase nevertheless shared with them a focus on the movement of resources in very small markets. The result was a new brand of institutionalism that was far more palatable to neoclassicists but largely unacceptable to traditional institutionalists. Today New Institutional Economics (NIE) occupies a prominent position within the mainstream of economic thought.

This revised brand of institutionalism has been a principal intellectual source of theory for modern law and economics. The problem with neoclassicism, outside from Pigou, is that it did not develop a coherent theory that the movement of resources from one place to another is costly, and that institutions are the devices that humans create in order to effect this movement. By contrast, old institutionalism understood the need for institutions well enough and thus was able to study the relationship between law and economics in a way that was previously unknown. Without marginalism as an analytic lever, however, old institutionalism was unable to do much more than catalog and classify. The merger of marginalism into institutionalism gave the new brand of institutionalism a much more coherent program and voice. However, old and new institutionalism are extraordinarily different creatures, with very different perspectives on the relationship between economics and law.

I. BRITISH MARGINALISM BEFORE COASE

Two characteristics of Cambridge economics proved to be lightning rods for Coase’s thought. The first was marginalism and the second was a belief that interpersonal comparisons of utility were a useful part of economic inquiry.

A. Marginalism, Equilibrium, and the Cost of Moving Resources

The word “marginalism,” with its conception of the rational actor and decision making at the margin, was very likely coined by John A. Hobson shortly before World War I to describe economics’ integration of marginal utility theory with classical political economy. But marginalist analysis dated back to at least the


By the early twentieth century, marginalist thought had overrun economics in Britain, the United States, and much of the Continent.24 Marginal utility theory revolutionized economics in a number of ways. First, it was forward looking. For the marginalists, value became a function of willingness to pay for the next unit, rather than an average of previous values or some theory about how much labor went into producing something, as it had been for the classical political economists.25 Marginalism’s rational actor received decreasing marginal value from goods and maximized his position by exchange until the marginal value of each was the same, or in “equilibrium.” In The Nature of the Firm, Coase simply applied this principle to the business firm rather than the individual as an economic actor.26

The early marginalists, particularly Alfred Marshall, were infatuated with mechanics, and with the ways that higher mathematics could be applied to questions about how physical materials tend to move toward a steady state in which all the forces acting on them are equalized.27 As in the theory of mechanics, marginalism introduced mathematics into economics in a way previously unknown in British economics, except for a few acknowledgements of the work of Augustin Cournot.28

At the same time, marginalism threatened to deprive economics of much of its historical, biological, and empirical content. Its leading thinkers were heavily preoccupied with developing the mathematical and geometric techniques of their discipline, and their principal focus was market behavior.29 Marginalism was quite unconcerned with how people actually experienced needs and wants, or what the inherited or environmental sources of wants might be. It simply presumed that its actors were rational, which meant that they had a set of transitive preferences30 over
goods that could be freely exchanged. When thinking of firms, the marginalists did
the same thing that they did for persons; that is, they were very attentive to the
mathematics of competition and monopoly, but paid little attention to the firm’s
internal operations or decision making. The best examplar of the tradition is Alfred
Marshall’s great \textit{Principles of Economics}, first published in 1890.\footnote{ALFRED MARSHALL, PRINCIPLES OF ECONOMICS (1890).} The
Marshallian business firm is largely a black box, consisting of cost functions and
responding to demand functions.\footnote{See Lionel Robbins, \textit{The Representative Firm}, 38 \textit{ECON. J.} 387, 389 (1928) (noting
theory is largely intended to ascertain its behavior as a demander of inputs and a supplier of
outputs); Alan J. Meese, \textit{Monopolization, Exclusion, and the Theory of the Firm}, 89 MINN. L. REV. 743 (2005) (similar).} As Coase himself observed many years later, the
firm in neoclassical economics was largely an abstraction that made its presence
known on the market as purchaser and seller, but whose internal workings were
[i.e., price theory] is to understand how the price system coordinates the use of resources, not
to understand the inner workings of real firms.”); Harold Demsetz, \textit{The Theory of the Firm Revisited}, 4 \textit{J.L. ECON. & ORG.} 141, 143 (1988) (noting extent to which the firm in
neoclassical economics is an abstraction).} In general, neoclassical price theory assumed a frictionless
system in which firms were rational maximizers constrained by competition.\footnote{See, e.g., FRANK H. KNIGHT, RISK, UNCERTAINTY, AND PROFIT 3–5 (1921) (noting
neoclassical economics’ highly abstract assumptions); GEORGE J. STIGLER, \textit{THE THEORY OF COMPETITIVE PRICE} 24 (1942) (same).}

Marginalist economics was heavily preoccupied with “equilibrium,” or the place
where resources ended up as people, firms, and markets reflected the assertion of
individual preferences.\footnote{See Blaug, supra note 10, at 388–92.} Prior to Pigou, however, the marginalists largely ignored the
\textit{cost} of moving resources from one place to another. They assumed that
individuals and firms arranged their own preferences and also that markets
reflected interpersonal trading. But in doing so, the marginalists presupposed that
all economic actors shift resources from one use to another without any frictions in
the system.\footnote{See DOUGLASS C. NORTH, \textit{STRUCTURE AND CHANGE IN ECONOMIC HISTORY} 5 (1981).

\[T\]his neoclassical formulation appears to beg all of the interesting questions.
The world with which it is concerned is a frictionless one in which institutions
do not exist and all change occurs through perfectly operating markets. In short,
the costs of acquiring information, uncertainty, and transactions costs do not
exist.} This lack of attention to the “costs of movement,” as Pigou called it,\footnote{See Pigou, supra note 4, pt. II, ch. III, § 3, at 138.} exacerbated the high level of abstraction in neoclassical economic reasoning.
Marshall himself was aware of the problem of abstraction in his *Principles*. In 1919, late in life, he published a much less theoretical and more descriptive study of business firms entitled *Industry and Trade*.\textsuperscript{38} Marshall’s principal argument was that as the geographic range of trade expands producers are in a better position to take advantage of economies of scale and scope, and thus plants and firms grew larger. This in turn shifted industries toward increased use of mass production. Marshall’s book contains detailed, descriptive discussions of the operations of firms in many markets, focusing heavily on products that had been the subject of fairly recent technological advances, such as steel and textiles.\textsuperscript{39}

Marshall did not develop a theory of business firm structure based on the cost of moving resources, or transaction costs. However, this passage from *Industry and Trade* suggests at least a minimal perception that transaction costs as well as production costs played a role in determining the vertical structure of the firm:

> The advantages, which a business derives from vertical expansion, are chiefly in regard to the organization of its work, and to the economies of marketing. As a rule it obtains few additional economies on the technical side of production: for there results little or no increase in the resources of plant and skill which it commands for any particular task. But it can so adjust the output of its lower stages to the requirements for material of the upper stages, that scarcely anything needs to be bought except for the lowest stage; and scarcely anything needs to be marketed except from the highest stages: while in the special case of a steel business, as has already been observed, it is often possible to shift temporarily some labour and some plant from work for one stage to work for another. In any case the higher stages can rely on the quality of the material supplied from the lower, and on its adaptability to their wants; especially if “planning” and “routing” are organized scientifically.\textsuperscript{40}

In *The Nature of the Firm* (1937), Coase spoke mainly of “marketing” costs rather than transaction costs.\textsuperscript{41} The Marshall passage also anticipated the view developed later in transaction-cost economics that transactional rather than

\begin{itemize}
  \item \textsuperscript{39} See, e.g., Marshall, *Industry and Trade*, supra note 38, at 56–59, 799–802 (discussing textile production and markets); id. at 94–95, 218–21, 802–03 (discussing steel production and markets).
  \item \textsuperscript{40} Id. at 589.
  \item \textsuperscript{41} Coase, *Nature of the Firm*, supra note 26, at 392 (“[T]he operation of a market costs something and by forming an organization and allowing some authority (an ‘entrepreneur’) to direct the resources, certain marketing costs are saved.”); see also id. at 395 (“[A] point must be reached where the loss through the waste of resources is equal to the marketing costs of the exchange transaction in the open market . . . .”).
\end{itemize}
technological constraints determine when a firm produces internally and when it procures from an external source.42

However, Marshall clearly did not share Coase’s idea of the firm. In another passage of Industry and Trade he opined that there was a tendency on the part of business firms to grow larger than was economically justified. He blamed this tendency on the aggrandizement of directors, suggesting that it showed up “in the tendency of some joint stock companies and municipalities to make things, which it would perhaps have been better for themselves and for others that they should have bought.”43 Marshall added that the “owner of a business, when contemplating any change, is led by his own interest to weigh the whole gain that it would probably bring to the business, against the whole loss.”44 But he concluded that the “private interest of the salaried manager, or official, often draws him in another direction.”45 These observations presaged not only Coase’s marginalist theory of firm structure, but also the idea of separation of ownership and control that became a staple of institutionalism’s economic critique of the business corporation in the 1930s.46 But neoclassical economics largely ignored Marshall’s observations and developed a theory of the firm that praised rather than lamented the separation of ownership and control, seeing the firm as an engine for the maximization of its own value quite apart from the separate wishes of stockholders.47

While the costs of moving resources from one place to another were not explicitly identified in Marshall’s model of markets, he was clearly aware of them. Indeed, Industry and Trade was obsessed with the notion that firms continuously seek to reduce their own costs. Marshall admired and often referred to “Taylorism,” or scientific efficiency analysis.48 Frederick Winslow Taylor was an American engineer who developed a theory of scientific management that emphasized the streamlining of steps in workflow in order to reduce costs.49 The Marshallian firm economized by taking advantage of the best technology and limiting the number of costly steps in the production process.

Arthur Cecil Pigou, who succeeded Marshall as professor of political economy at Cambridge in 1908,50 developed a much more comprehensive theory about the

43. MARSHALL, INDUSTRY AND TRADE, supra note 38, at 322.
44. Id. at 324.
45. Id.
50. On Pigou’s appointment, see RONALD H. COASE, The Appointment of Pigou as Marshall’s Successor, in RONALD H. COASE, ESSAYS ON ECONOMICS AND ECONOMISTS 151,
costs of moving resources from places of lesser value to places of greater value. Pigou was also quite clear that the costs of bargaining were included among these costs of movement. As he wrote in The Economics of Welfare:

Suppose that between two points A and B the movement of a unit of resources can be effected at a capital cost equivalent to an annual charge of \( n \) shillings for every year during which a unit that is moved continues in productive work in its new home. In these circumstances the national dividend will be increased by the movement of resources from A to B, so long as the annual value of the marginal social net product at B exceeds that at A by more than \( n \) shillings; and it will be injured by any movement of resources which occurs after the excess of the value of the marginal social net product at B has been reduced below \( n \) shillings.\(^{51}\)

From this observation Pigou derived this rule:

[W]hen the costs of movement between A and B are equivalent to \( n \) shillings, the national dividend is best served by the maintenance of the existing distribution, whatever that may be, provided that this distribution does not involve a divergence in the values of marginal social net products greater than \( n \) shillings; and, if the existing distribution does involve a divergence greater than \( n \) shillings, by a new distribution brought about by the transference of sufficient resources to bring the divergence down to \( n \) shillings.\(^{52}\)

For Pigou, the cost of moving resources included bargaining costs, but was not limited to them. Rather, these costs included:

[P]ayments that have to be made to various agents in the capital market, promoters, financing syndicates, investment trusts, solicitors, bankers, and others, who, in varying degrees according to the nature of the investment concerned, help in the work of transporting capital from its places of origin to its places of employment.\(^{53}\)

Like the other early Cambridge economists, Pigou believed that the sovereign could increase welfare by transferring resources among classes, provided that the recipient groups experienced greater welfare than the expropriated groups gave up, and that the cost of movement was less than the difference. That is, Pigou believed

---


53. Id. pt. II, ch. VII, § 1, at 158. Guido Calabresi, the other person that Richard Posner cites as a founder of modern law and economics in the United States, also used “cost” much more broadly in this fashion. See Calabresi, supra note 1 (speaking not of transaction costs, but rather of such costs as those of running the insurance system, costs of administering the system of enterprise liability, and the like). For Posner's comment crediting Calabresi, see supra note 1.
that at least to some extent the economic policy maker could engage in interpersonal comparisons of utility.

B. The Ordinalist Revolution

At least as subsequent critics characterized it, the second important element of Cambridge economics was its belief that utility functions could be expressed in cardinal units that could be compared across persons.54 Each individual maximized personal utility by prioritizing desire and expenditure. But the Cambridge economists agreed that, at least at a general level, it was possible to compare utilities across people or classes of people—for example, by saying that the transfer of wealth from the wealthy to the poor increased social welfare because a poor person valued an incremental (marginal) dollar by a greater amount than the wealthy person suffered from its loss.

This view was rarely expressed as categorically as Robbins later implied in his Essay.55 Even early British marginalists, such as William Stanley Jevons and F.Y. Edgeworth, expressed some doubt about the domain of utility measurement. Jevons believed that comparison of subjective mental states was impossible, but that welfare comparisons across social classes based on externally observable criteria might be feasible.56 In his Mathematical Psychics, Edgeworth suggested that a device called a “hedonimeter” might someday be developed that could measure individual utilities and even compare the utility of one person with that of another.57 Italian economist Vilfredo Pareto, writing in the 1890s, argued that the satisfaction of desires comes in two kinds. One refers to the satisfaction of subjective preferences, “whether legitimate or not.”58 Pareto gave the term

55. Lionel Robbins, An Essay on the Nature & Significance of Economic Science 137 (2d ed. 1935) (“It is safe to say that the great majority of English economists accept them as axiomatic.”).
56. See William Stanley Jevons, The Theory of Political Economy (Penguin Books 1970) (1871). In the introduction, Jevons stated: The reader will find, again, that there is never, in any single instance, an attempt to compare the amount of feeling in one mind with that in another. I see no means by which such comparison can be accomplished. The susceptibility of one mind may, for what we know, be a thousand times greater than that of another. But, provided that the susceptibility was different in a like ratio in all directions, we should never be able to discover the difference. Every mind is thus inscrutable to every other mind, and no common denominator of feeling seems to be possible. . . . [T]he motive in one mind is weighed only against other motives in the same mind, never against the motives in other minds.
58. Cooter & Rappoport, supra note 54, at 515 (quoting Vilfredo Pareto, Cours d’Economie Politique 3 (1896)).
“ophelimity” to these purely subjective preferences. He reserved the term “utility” to describe the second kind of desires—those which are “conducive to the development and prosperity of an individual, a people, or the human race.” In sum, for these early marginalists, utility had not only a subjective meaning, but also an objective meaning that was related to observed productivity. By and large they doubted their ability to make interpersonal comparisons of subjective preference, but had greater confidence about interpersonal comparisons of these objective criteria of welfare.

Both Marshall and Pigou believed that human utility functions were more-or-less similar across persons, and so it followed that interpersonal comparisons of cardinal utilities were at least roughly possible. As a result, both Marshall and Pigou could believe that classwide involuntary wealth transfers from the wealthy members of society to poorer members increased total welfare. Having less to start with, poorer individuals experienced higher marginal utility from an increment of money than did the wealthy. For example, Marshall wrote in 1885 that social change could beneficially “tak[e] account of the fact that the same sum of money measures a greater pleasure for the poor than for the rich.”

Pigou also believed that wealth transfers from wealthy to poor typically took luxuries away from the wealthy in order to provide necessities to the poor:

[I]t is evident that any transference of income from a relatively rich man to a relatively poor man of similar temperament, since it enables more intense wants to be satisfied at the expense of less intense wants, must increase the aggregate sum of satisfaction. The old “law of diminishing utility” thus leads securely to the proposition: Any cause which increases the absolute share of real income in the hands of the poor, provided that it does not lead to a contraction in the size of the national dividend from any point of view, will, in general, increase economic welfare.

Pigou’s position was in fact quite measured. More than Marshall, Pigou wanted to preserve the status of economics as a tool for policy. He realized that doing so forced economics to confront moral issues and perhaps even step over “the narrow boundaries of . . . science” and into ethics. Whatever one might think of the ordinalist critique of interpersonal utility comparisons, Pigou wrote in the early 1950s, “[i]n all practical affairs we act on” the supposition that such comparisons

59.  Id.
60.  Id.
63.  PIGOU, supra note 4, pt. I, ch. VIII, § 3, at 89.
are meaningful, and that without them the entire “apparatus of practical thought,” would be ruined.\textsuperscript{65}

Further, Pigou believed that his arguments for wealth transfers from the wealthy to the poor did not rest on any notions about utility at the margin, but rather on the potential for creating objectively measurable improvements. Such transfers to the poor “make possible the development in them, through education and otherwise, of capacities and faculties adapted for the enjoyment of the enlarged income.”\textsuperscript{66} This idea created a decisive line between the neoclassicists and the institutionalists. For Robbins, human welfare was measured purely as a consumption function, with utility as its currency. For Darwinians and the emergent social sciences, the human welfare function came closer to resembling an economists’ \textit{production} function, which evaluates an entity by looking at its output in relation to its inputs. That is, it tended to equate welfare with the ability to produce rather than with the satisfactions that come from consumption.\textsuperscript{67}

As noted, one important qualifier that the Cambridge economists placed on interpersonal utility comparisons is that they were thought to be more meaningful, or robust, when applied to groups rather than individuals. Further, this idea made economics more relevant to policy, because government wealth transfers typically worked among classes or identified homogeneous populations rather than individuals. In one well-known passage of his \textit{Principles}, Marshall conceded that one would not be warranted in saying that two \textit{individuals} of equal wealth derived the same marginal utility from a fixed amount of wealth.\textsuperscript{68} However,

\begin{quote}
\textit{[i]f there are a thousand persons living in Sheffield, and another thousand in Leeds, each with about £100 a-year, and a tax of £1 is levied on all of them; we may be sure that the loss of pleasure or other injury which the tax will cause in Sheffield is of about equal importance with that which it will cause in Leeds . . . .}\textsuperscript{69}
\end{quote}

Pigou likewise spoke mainly of wealth transfers among classes. “[U]nless we have a special reason to believe the contrary, a given amount of stuff may be presumed to yield a similar amount of satisfaction, not indeed as between any one man and any other, but as between representative members of groups of individuals . . . .”\textsuperscript{70}

For these economists, sometimes referred to as the “material welfare” school, interpersonal comparisons of utility were possible at least at the “survival” or “basic necessity” end of the utility range—that is, for those goods that people needed first and could not afford to be without. However, as one proceeded along

\begin{itemize}
\item \textsuperscript{66} Pigou, \textit{supra} note 4, pt. I, ch. VIII, § 3, at 91.
\item \textsuperscript{67} For a comprehensive attempt to relate human “utility functions” to Darwin and survival, see Richard Dawkins, \textit{River out of Eden: A Darwinian View of Life} 95–134 (1995).
\item \textsuperscript{69} Id. at 18–19.
\item \textsuperscript{70} Pigou, \textit{supra} note 65, at 292 (emphasis in original).
\end{itemize}
the marginal utility scale to “comforts” and “luxuries,”71 such comparisons were much more difficult to undertake.72

Even highly mathematical pre-ordinalist economists, such as Yale’s Irving Fisher, developed statistical concepts of the “average family” in order to measure the marginal utility of money.73 He concluded that “mass statistical measurements” would be more accurate than the measurement of the utility of “any individual who feels them.”74 From this Fisher devised what he believed were average utility levels for different economic classes, which he believed justified a progressive income tax. Several American economists developed similar ideas.75

Fisher’s concern with statistical averages illustrates the one important weakness that members of the material welfare school perceived about interpersonal utility comparisons—the great difficulty of measurement. Marginalists since the time of Jevons had been able to apply mathematics to problems of individual choice in

71. Marshall, Principles of Economics, supra note 68, at bk. II, ch. iii, § 3, at 67–68 (“This brings us to consider the term Necessaries. It is common to distinguish necessaries, comforts, and luxuries; the first class including all things required to meet wants which must be satisfied, while the latter consist of things that meet wants of a less urgent character. But here again there is a troublesome ambiguity. When we say that a want must be satisfied, what are the consequences which we have in view if it is not satisfied? Do they include death? Or do they extend only to the loss of strength and vigour? In other words, are necessaries the things which are necessary for life, or those which are necessary for efficiency?”).

72. The distinction is developed in Cooter & Rappoport, supra note 54, at 513–14.


75. See, e.g., John R. Commons, The Distribution of Wealth 10 (1893) (“To the poor man all the marginal increments may afford high satisfaction, because his supplies are limited; but to the rich man the marginal increments may give little satisfaction. Yet each, though on different levels, endeavours to make the marginal increments in all lines equal.”); 1 F.W. Taussig, Principles of Economics ch. 9, § 6, at 132 (3d ed. 1921) (“The principle of diminishing utility, if applied unflinchingly, leads to the conclusion that inequality of incomes brings a less sum of human well-being than equality of incomes, and that the greater the inequality, the less the approach to the maximum.”); George P. Watkins, Welfare as an Economic Quantity 185–90 (1915) (arguing that total social utility would increase if incomes were leveled somewhat and the right to inherit severely curtailed); J.B. Clark, The Ultimate Standard of Value, 1 Yale Rev. 258 (1892) (arguing that social welfare is a function of aggregated individual utilities); Simon N. Patten, The Scope of Political Economy, 2 Yale Rev. 264 (1893) (arguing that homogeneity of wealth and property tends to increase total utility); Jacob Viner, The Utility Concept in Value Theory and Its Critics (pts. 1 & 2), 33 J. Pol. Econ. 369, 638, 644 (1925) (“Changes in the relative distribution of income as between different classes will bring about changes in the amount of welfare, even though the aggregate real income of the community remains the same.”).
technically and conceptually accurate ways. Tautologically, an individual equates utilities at the margin—exchanging goods or time for which he has lower utility for that which provides higher utility—until all are in equilibrium. But neoclassicism became stuck at that point. If utilities could not be compared across persons, then economics could say little to nothing about the optimal distribution of wealth and, indeed, could not even justify providing food to starving children.

The Robbins critique of interpersonal utility comparisons was forceful not simply because it pointed out the greater susceptibility of individual preferences to mathematical reasoning. Rather, it defined interpersonal utility comparisons as being outside the boundaries of economic science altogether. This well-known paragraph from Robbins’ *Essay* on economic methodology states his conclusions:

> But suppose that we differed about the satisfaction derived by A from an income of £1,000, and the satisfaction derived by B from an income of twice that magnitude. Asking them would provide no solution. Supposing they differed. 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 discover what is going on in B’s mind, nor B to discover what is going on in A’s. There is no way of comparing the satisfactions of different people.76

And speaking of comparing one person’s utilities with those of another:

> It is a comparison which necessarily falls outside the scope of any positive science. To state that A’s preference stands above B’s in order of importance is entirely different from stating that A prefers $n$ to $m$ and B prefers $n$ and $m$ in a different order. It involves an element of conventional valuation. Hence it is essentially normative. It has no place in pure science.77

As a result, Robbins concluded that any argument for greater income equality based on the theory that the poor would derive greater marginal utility from an increment than the wealthy would lose from an identical reduction is “entirely unwarranted by any doctrine of scientific economics.”78 Rather, “[i]t rests upon an extension of the Law of Diminishing Marginal Utility into a field in which it is entirely illegitimate.”79

It is worth noting that there were and remain good alternatives to the Robbins approach of writing all interpersonal utility comparisons out of the discipline. Most obviously, Robbins’s critique was expressly about individual subjective mental

76.  ROBBINS, supra note 55, at 139–40 (emphasis in original); see also Lionel Robbins, *Interpersonal Comparisons of Utility: A Comment*, 48 ECON. J. 635 (1938).
77.  ROBBINS, supra note 55, at 139.
78.  Id. at 137.
79.  Id.
states—something that earlier marginalists had already decided was beyond measurement. By looking for alternative, more objective criteria of welfare, neoclassicism could have developed separate branches with their own individual methodologies for price theory and welfare economics. Ten years before Robbins’s Essay, University of Chicago economist Jacob Viner faulted marginalist economics for purporting to measure individual subjective welfare but instead using an objective standard for everything except voluntary exchanges. Viner wrote, “[M]uch of what passes for utility theory is really objective price-theory presented in the purloined terminology of subjective analysis . . . .”81 Frank Knight criticized the “pernicious concept of utility.”82 He allowed that “[w]hen we come to pass judgment on the workings of the price system, we have to have a theory of utility as a starting point.”83 In Knight’s concept, a purely ordinalist notion of declining marginal utility for the individual helped us explain the workings of markets. However, it was useless as an explanation of social behavior or social welfare.

More importantly, Robbins’s conclusions about the domain of economics were just as normative as those of the Cambridge economists. The idea that making welfare judgments outside of the limitations of subjective utility analysis was “normative”—while strict adherence to the non-comparability of individual utilities was somehow scientific—itself reflected a highly value-laden judgment about the scope of scientific investigation. If one views welfare as nothing more than subjective preference, then interpersonal comparisons cannot be observed. As a result, they must be counted as normative under a criterion for science requiring that conclusions be capable of being falsified. On the other hand, if one views individual welfare as a function of productivity rather than subjective desire, then empirical comparisons are possible. For example, we could conduct observations that $1000 given to a poor person made noticeable differences in nutrition, housing, education, and the like, while $1000 given to a millionaire made no measurable difference at all. Cambridge economists such as Pigou often measured welfare in this way, and both the institutionalists and American legal realists continued to do so in the 1920s and after. Indeed, Joan Robinson’s later work on the production function was based in substantial part on the premise that the value of individual human productivity had been slighted in welfare analysis.84

Even the most virulent ordinalists reverted to objective measures when they spoke of business firms rather than persons. Firms maximize value, which can be cardinaly measured and is directly related to productivity. The only way we can

81. Viner, supra note 75, at 657.
84. See Joan Robinson, The Production Function and the Theory of Capital, 21 REV. ECON. STUD. 81 (1953); see also Joan Robinson, The Production Function, 65 ECON. J. 67, 67 (1955) (using illustration of the production function of a robin as measured by the number of grubs caught in relation to the number of robin-minutes worked).
speak of value is objectively. Indeed, neoclassical economics rarely considers the states of mind of a firm’s shareholders or managers. So, for example, to observe that a firm turned inputs that cost $1000 into output worth $1200 could be a strictly scientific proposition; however, to attach welfare significance to the observation that a biological person responded to a $1000 gift by buying a better house or improving her education lays outside the boundaries of science because it was thought to depend on unverifiable inferences about state of mind.

In sum, Robbins’s critique rested on the premise that it was “scientific” to speak economically about human beings in terms of consumption but “normative” to speak of them in terms of production. This constraint is itself normative, but it is also irrational because greater productivity reduces scarcity and thus leads to greater consumption as well. One does not need to engage in interpersonal comparisons of utility to conclude that \textit{ceteris paribus} a society that has greater wealth will permit a greater amount of wants to be satisfied.

Robbins’s limitations were particularly stringent because his \textit{Essay} did not purport to be simply a set of constraints on the study of personal preferences. Rather, he represented that he spoke about the entire field of economics as a science. In a few paragraphs, he identified human welfare with consumption rather than production and proclaimed the study of anything else as purely normative. As developed below, one impact of this limitation was to divorce neoclassical economics from the theory of evolution and evolutionary psychology, the alternative claimants to scientific status in the late nineteenth and early twentieth century. At bottom, Darwinian evolution is a theory about production—that is, about the development of traits that enable organisms to survive in an environment of scarce resources. This became clear with the rise of Reform Darwinism prior to the Progressive Era. Metaphors typically used by the Reform Darwinists included the farmer cultivating seed or the cattle breeder; these were interventions in nature whose efficacy depended on measured changes in productivity, not on subjective feelings of well being. Success in the Darwinian struggle for survival is plotted as a production function, not a consumption function.

Robbins’s critique of interpersonal utility comparisons also gave insufficient weight to the enormous trade-offs that had to be given up. As Cooter and Rappoport have concluded, any evaluation of the ordinalist critique of interpersonal utility comparisons makes it “necessary to balance the gains in understanding markets which the ordinalist framework facilitated against the losses in understanding human welfare, suffered by abandoning the material welfare framework.” Nevertheless, as a matter of history, one thing seems clear: after the ordinalist revolution swept the field, economic science was all about voluntary exchange, or markets, and markets are all about transactions.

Finally, the ordinalist critique revealed that, depending on assumptions, the mathematics of marginalism could lead to ideological views at the two extremes...
but with a middle ground that was difficult to defend. Carried to its logical conclusion, the principle of diminishing marginal utility, coupled with interpersonal utility comparisons on an individual basis, led directly to socialism, or a state in which wealth was equalized, for equalization produced the greatest total welfare. Both Alfred Marshall and particularly Joan Robinson, pre-ordinalists, became near socialists late in their lives—a fact that may have operated to deny the ageing Joan Robinson the status of first woman to receive the Nobel Prize in Economics.89

At the other extreme, the Robbins critique saw only individual orderings of preferences and market exchange as increasing welfare, and proclaimed that drawing conclusions about the welfare effects of involuntary wealth transfers lay outside the boundaries of economic science. The result of that view was extreme laissez faire, with a government justified in intervening on welfare grounds only in infrequent cases of market failure.90

Compromise positions were difficult to defend. The proposals of people such as Marshall, Pigou, and Irving Fisher that interpersonal utility comparisons be based on classes rather than individuals was a well-intentioned attempt. But it ultimately did not fit into either paradigm particularly well. For the older marginalists, the attempt at compromise led to equality of income between classes rather than individuals. But once again the economic effect was pure socialism. For the ordinalists, class comparison was no more readily susceptible to scientific study than individual comparison. So economics was at an impasse.

While Coase never engaged in the ordinalist debate and was seriously interested in institutions, he was also completely ordinalist in his thinking. As a result, for him the concern of economics was properly limited to three things: (1) market exchange and the costs of the exchange process; (2) price theory, which employed a currency of constant value; and (3) the internal preferences exercised by the single economic actor, including the business firm, making decisions under scarcity.

II. THE RISE OF INSTITUTIONALISM

A. Darwin, Behaviorism, and the Revolt Against Marginalism

The two greatest intellectual products of Victorian thought were Darwinism and marginalism.91 While Darwin’s theory of evolution by natural selection was uniquely British, marginalism evolved simultaneously in both England and the Continent. Its principal innovators were William Stanley Jevons of England, Carl Menger of Austria, and Leon Walras of France.92 Both Darwinism and marginalism
had effects that not only overturned their “home” disciplines but that also spilled out into many other areas. During the final decades of the nineteenth century and the first two decades of the twentieth, Darwinism and marginalism seemed independently to paint a fairly complete picture of human welfare and social obligation. Marginalism provided a rigorous theory of individual incentives and valuation. Darwinism offered both a principle of physical development and also a theory of society and culture, as well as prescriptions about welfare and the role of the state.

One characteristic of Darwinism was the link it saw between instinct and survival, and the fact that within a species basic survival needs were nearly identical. In any given environment, the organisms that tended to survive had a common set of characteristics and needs. These largely biological observations led people in diverse ideological directions, at least as caricatured in later writing by intellectual historians. For so-called “Social Darwinists,” the fittest organisms will survive and state interference to support them artificially will degrade the race or retard its evolution. At the other extreme, the “Reform Darwinists” believed that both the environment and the individual could and should be manipulated in order to make individuals better able to survive and prosper. For the Reform Darwinists, as well as most of the emergent generation of social scientists, Darwinian thought provided a basis for thinking that by altering the environment, Darwinism and marginalism seemed independently to paint a fairly complete picture of human welfare and social obligation. Marginalism provided a rigorous theory of individual incentives and valuation. Darwinism offered both a principle of physical development and also a theory of society and culture, as well as prescriptions about welfare and the role of the state.93


95. See Hofstadter, supra note 94, at 6–7.

or giving its members better survival tools, they could increase their chances of success. Darwinism thus permitted liberal policy makers to generalize about individual “welfare” needs without making assumptions about individual mental states. As a result, the ordinalist economic critique of interpersonal utility comparisons was irrelevant to them.97

Among the literate laity, Darwinism was much better known than marginalism but also more infamous. Marginalism never produced any equivalent to Richard Hofstadter’s extremely influential book entitled Social Darwinism in American Thought,98 as well as an entire generation of historians who explained intellectual history largely in Darwinian terms.99 Darwinism was better known because, at a fairly general level, one could explain the theory of natural selection without the use of technical apparatus such as mathematics. Darwinism was also more infamous, however, because the “natural selection” that motivated Darwinism was more threatening to Christianity than the “hedonism” that permeated the marginalist’s subjective value system. Orthodox pastors and priests never railed at marginalism the way they did at Darwinism. As a result, epithets such as “Social Darwinism” were used in public discourse to describe essentially economic thinking that was in fact more properly attributable to neoclassical economics and marginalism than to the theory of evolution by natural selection.100

Robert Cooter and Peter Rappoport’s provocative article on the ordinalist revolution compares the “material welfare” school, which believed that limited interpersonal utility comparisons were possible, with the ordinalists, who soundly rejected their views and virtually wrote such comparisons out of the boundaries of economic science. However, Cooter and Rappoport never mention Darwin or the theory of evolution.101 Nevertheless, built into Darwinism was a very powerful


98. HOFFSTADTER, supra note 94.


101. See Cooter & Rappoport, supra note 54. The only allusion to the theory of evolution occurs in a footnote discussing Pareto that refers to his conception of material welfare as “some Spencerian ideas of evolutionary fitness.” Id. at 516 n.22.
theory of “preference” that is quite different from the radically individualistic and subjective “preference” that is the topic of marginalist economics. This Darwinism strongly affected many of the first generation of economists that we today describe as institutionalists, although in different ways.  

For marginalists in the neoclassical tradition, preferences were considered as naked, which means that they were simply recognized as asserted with little or no thought about their derivation. Willingness to pay is what it is, and querying why someone prefers this or that is simply outside the boundaries of economic science. For Robbins, whose thought revolutionized the discipline, the concept of scarcity meant little more than the fact that a person cannot have all that she desires of everything, so choices and rankings are essential. As a result, the combination of scarcity and preference give rise to the concept of declining marginal utility. An individual maximizes total utility by acquiring a blend of everything she values until, at the margin, she values everything that she has by the same amount. For Robbins, this required only that an individual have a set of choices, that this set be rational (i.e., transitive), and that no one can have absolutely all they might want of everything.

In sharp contrast, Darwinism’s starting point for analysis was the environment rather than the individual. Ironically, Robbins’s scarcity definition of economics that became so influential in the development of the new welfare economics had in fact been a staple of Darwinian thought for decades. Indeed, pre-Darwinians, such as the classical political economist Thomas Malthus, had incorporated it into

103. ROBBINS, supra note 55, at 78–79 (“The main postulate of the theory of value is the fact that individuals can arrange their preferences in an order, and in fact do so.”).
104. See id. at 91–92 (“[I]n a state of equilibrium the relative significance of divisible commodities is equal to their price does involve the assumption that each final choice is consistent with every other, in the sense that if I prefer A to B and B to C, I also prefer A to C . . . .”).
105. See Robbins’s famous definition of economics: “Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.” Id. at 16.
his theory of population, and Darwin, who was strongly influenced by Malthus, built the theory of natural selection on the principle of scarcity.108

The most important difference between Darwinian choice under scarcity and Robbins’s definition had to do with the role of natural selection, which in turn controlled the degree of individual autonomy that accounts for the choice mechanism.109 For Robbins’s definition, the individual was absolutely autonomous in the sense that one cannot scientifically get behind the choices; they simply are what they are and economic science deals with the way individuals can (positive) and should (normative) act so as to maximize their satisfactions.

For Darwinians, in very sharp contrast, choice is a fundamental part of one’s survival mechanism. This view of preference has powerful implications for theories of both individual and social choice. First, under the Darwinian view, individuals do not “control” their choices, certainly not in the broad unconstrained sense that neoclassical economics has come to adopt. Second, Darwinian natural selection entails that over a wide range the choices individuals make—and certainly individuals in the same species and environment—are identical or at least closely similar. Behaviorism increasingly came to see choices as “conditioned,” or as a consequence of repeated iterations of some motion or response that produced either pleasure or pain. Third, the Darwinian perspective dictates an externally imposed preference ordering: namely, survival first and “surplus” later. For many organisms in nature, there is no surplus and survival dictates virtually every choice that is made. For human beings in modern society, the story is more complex. Some still have very little surplus while others have a great deal. But the all-important fact for Darwinism, although not stated in this terminology, was that in the survival range organisms have preference orderings that are closely similar and that could be treated as identical for policy purposes. Indeed, even the conditions that Robbins imposed on choice as inherent to rationality—namely, equation of preferences at the margin and transitivity—are “survival” traits, in the sense that organisms that do not tend to their needs in an economical way lose out in the evolutionary struggle.

Darwinian social theorists had made these observations throughout the last half of the nineteenth century, and they formed the basis for the sharp divide between


so-called “Social” and “Reform” Darwinists. Social Darwinists believed that the struggle for survival and the emergence of “superior” characteristics were the keys to human progress. As a result, the best policy that the State could pursue was to leave the struggling to themselves, without aid, letting nature determine who was fit to survive.

By contrast, Reform Darwinists, the forerunners of American Progressivism, began by contrasting the fate of the organism in nature with the organism under domestication. They observed that human beings are unique among organisms in that they are aware of the fact that they are evolving and in a position to control the process. For example, only human beings are involved in domestication, or the raising, breeding, and improvement of other organisms by management of the environment—a process that need be no more complicated than farming or cattle breeding.

Most important for present purposes was the Reform Darwinists’ attitude toward scarcity. Social Darwinists such as Herbert Spencer tended to see the resources produced by nature as fixed, and an excess of organisms including human beings as competing for them. In contrast, Reform Darwinists emphasized humankind’s ability to increase the amount of resources by vast amounts, perhaps even turning shortages into surplus. For example, Lester Ward, who was virtually deified by some liberal historians in the 1950s, argued that Social Darwinism saw “all the functions of society” as “performed in a sort of random, chance manner, which is precisely the reverse of economical, but wholly analogous to the natural processes of the lower organic world.” Ward contrasted the wasteful “biological economy” of nature with the “psychological economy” of “rational man.” The rise of agriculture was a clear example. Nature sows by leaving seed “to the wind, the

110. See Hovenkamp, Evolutionary Models, supra note 93, at 655–56.
113. See Hofstadter, supra note 94, at 38–44 (discussing Spencer’s beliefs).
115. 3 Ward, Glimpses of the Cosmos, supra note 112, at 35; see also 1 Ward, Dynamic Sociology, supra note 112, at 542.
water, the birds and animals.”117 In contrast, the human farmer “prepares the ground, clearing it of vegetable competitors, then . . . carefully plants the seeds at the proper intervals . . . and after they have sprouted . . . keeps off their enemies, . . . supplies water if needed, and even supplies the lacking chemical constituents of the soil . . . .”118 Institutionalist economist John R. Commons echoed this theme as well, contrasting the “natural selection” of Darwinian nature with human “artificial” selection, which involves “intention, and belongs to human reason.”119

Reform Darwinism was the guiding principle of the early pioneers of the modern social science movement.120 The result of their thinking was behaviorism—a methodology that ordinalists such as Robbins flatly rejected.121 Of course, every methodology must draw inferences from what is observed. As Robbins conceded, economists cannot observe “valuation” because it is entirely “a subjective process.”122 However, we can observe willingness to pay as measured by actual transactions or bids. Behavioral social science could not observe valuation either, but it could observe other external manifestations of welfare, such as physical health, productivity, education, savings, and the like. As Reform Darwinists such as Ward observed, human managers relied on external, or objective, indicia of welfare in their domestication of lower organisms from plants to cattle.

B. Institutionalist Economics from Veblen to the Legal Realists

Mark Blaug sums up institutionalism in a single sentence, as an attempt “to persuade economists to base their theories, not on analogies from mechanics, but on analogies from biology and jurisprudence.”123 Institutionalism as a recognized movement began during World War I when a group of economists, led by John

117. Id. at 355.
118. Id.
121. See ROBBINS, supra note 55, at 87–88 (addressing critique that preferences cannot be observed, but that behavior can be, so behaviorism would provide a better tool for economic analysis).
122. Id. at 87.
123. BLAUG, supra note 10, at 700.
Maurice Clark, Wesley Clair Mitchell, Walton Hamilton, Rex Tugwell, and a little later John R. Commons and Morris Copeland, identified a set of economic positions that distinguished their views from those of the prevailing marginalism.124

Their acknowledged champion was Thorstein Veblen, whose productive period was largely over by the early 1920s. Veblen spent much of his career attacking marginalism and neoclassicism for not taking Darwinian theory into account. Veblen was the most explicitly Darwinian of the major institutionalists, emphasizing the role of survival instincts in forming individual preferences as well as social institutions themselves.125 Veblen thought of human choice in terms of “habits” rather than preferences, and habits were entirely the product of evolution. Within this model, Veblen believed that marginalist theory was positively harmful and that marginalism’s continual search for equilibrium was misguided, given that evolution knows no equilibrium but is continuously rematching environment and the organisms contained in it.126 For Veblen, institutions were human creations for directing wealth, power, and social control. They were entirely a product of evolution, and they were worthy of economic study because they tended to be of much longer duration than the individuals who lived in them. An institution was “a usage which has become axiomatic and indispensable by habituation and general acceptance.”127

While all of the institutionalists were influenced by the theory of evolution and the emergent American social sciences, explicitly Darwinian rhetoric was less prominent in the writings of the inter-war institutionalists than it had been in Veblen. By and large, however, the institutionalists were evolutionists in their biology and increasingly after 1920, behaviorists in their psychology. For example, Hamilton emphasized the importance of using social psychology rather than

125. For a good analysis, see Hodgson, supra note 10.
127. Veblen, Absentee Ownership, supra note 126, at 101 n.1.
marginalist theory in order to understand human economic behavior.\footnote{128} Lawrence Kelso Frank, a Columbia-trained economist who spent much of his career at the Rockefeller Foundation, made a powerful plea for more behaviorist theory in economics.\footnote{129} Wesley Mitchell, who served as president of both the American Economics Association and the American Statistical Association and later on became a major theoretician of social planning during the New Deal, called repeatedly for a more behavioral approach to economics.\footnote{130}

One theme that dominates much of the history writing about the institutionalists is that behaviorism led to their downfall.\footnote{131} As a principle of human action, behaviorism was sharply in conflict with marginalism because it was hostile toward any concept that involved a hypothesis about state of mind, including valuation. The behaviorists did not speak about anything as rational as valuation at the margin, but rather about “impulse,” and the origins of these behaviors lay deep in biological development and the struggle for survival. Marginalism as a value theory seemed uniquely descriptive of human beings as opposed to lower animals; but the behaviorists believed that thought was no different from a body’s other physical processes and that, in this regard, all organisms are alike.\footnote{132} Further, the scientist can observe only the behaviors, not any rational value scheme that might serve to rationalize them.\footnote{133} These views created an uneasy relationship at best with a profession obsessed with the elegant mathematics that marginalist analysis was producing. In sharp contrast, behaviorist social science in the 1920s and 1930s was highly empirical and never developed a unifying theory with predictive power. The result was empirical studies that aggregated enormous amounts of factual matter but without important, testable unifying conclusions.\footnote{134} Coase himself would later

\footnote{128.  See Hamilton, supra note 124, at 316.}
\footnote{129.  See Lawrence K. Frank, The Emancipation of Economics, 14 AM. ECON. REV. 17 (1924).}
\footnote{130.  See Mitchell’s collected essays in Wesley Clair Mitchell, The Backward Art of Spending Money (1937).}
\footnote{132.  See, e.g., John Dewey, Human Nature and Conduct: An Introduction to Social Psychology 106–09 (1922).}
\footnote{133.  See, e.g., Pier Francesco Asso & Luca Fiorito, Waging War Against Mechanical Man: The Knight-Copeland Controversy over Behaviorism in Economics, in 21-A Research in the History of Economic Thought and Methodology 65 (Warren J. Samuels & Jeff E. Biddle eds., 2003); Morris A. Copeland, Economic Theory and the Natural Science Point of View, 21 AM. ECON. REV. 67 (1931). University of Chicago economist Frank Knight, who toyed with various aspects of institutionalism throughout his career, was very critical of behaviorism because its list of “instincts” defied rigorous analysis. See Frank H. Knight, Ethics and the Economic Interpretation, 36 Q.J. ECON. 454, 467 (1922) (“If instincts are to be scientifically useful, it must surely be possible to get some idea of their number and identity. But there has always been substantially unanimous disagreement on this point. Logically the choice seems to lie between a meaningless single instinct to do things-in-general and the equally meaningless hypothesis of a separate instinct for every possible act.”).}
\footnote{134.  On behaviorism in American institutionalism, see generally Shira B. Lewin,
describe the work of the institutionalists as “not theoretical but anti-theoretical.” 135

The same can be said of the legal realists, who were the legal counterpart of institutionalism and who together produced an impressive mass of interdisciplinary empirical scholarship. 136 Robert L. Hale, an institutionalist economist who wrote mainly on legal issues, was among the first American academics to have a joint appointment at the law school and economics department at Columbia University. 137

One characteristic of institutionalism, both new and old, is that by broadening the reach of economic analysis beyond traditional markets, it is able to capture a more complete set of the mechanisms by which resources are moved from one place to another. As Karl Polanyi, who was sympathetic with American institutionalism, so powerfully observed in 1944, markets composed of traders and voluntary exchange have accounted for only a small fraction of the movement of resources throughout history. 138 Families, tribes, governments, firms, gift giving,
and a miscellany of other institutions have been responsible for most of it. Further, these alternative forms have often either reflected or produced much greater amounts of social solidarity than has the highly individualistic theory of competition underlying market exchange.

The historical and economic literature on American institutionalism is enormous, and this discussion offers only a brief description of this diverse group of economists, who did most of their work between World War I and the onset of the Great Depression. With some exceptions, they shared severe doubts about the efficacy of marginalism as a unifying explanatory model for economics. As a result, they largely rejected or ignored ordinalism, which is entirely derivative of marginalism. The institutionalist objection was not that they thought that interpersonal comparisons of utility were possible. Much more fundamentally, they disputed the premise that marginal utility defines economic welfare at all. Further, while neoclassicism tended to emphasize the radically individualistic and idiosyncratic nature of human preference, the behaviorist psychology that institutionalists employed emphasized its similarity. Organisms of the same species thrust into the same environment tend to respond in the same way.

The relation between institutionalism and Darwinism was complex and evolved from Veblen’s explicitly evolutionary rhetoric at the beginning of the century to a more empirical behaviorism in the 1920s. Most importantly, as their name implies, they believed that “institutions” deserve more explicit recognition in economic analysis than neoclassical economics had given them. Significantly, the institutionalists did not all agree with each other even on these fundamental tenets. For example, Thorstein Veblen was extremely hostile toward marginalist analysis, while John Maurice Clark believed it was essential. John R. Commons understood marginalist analysis and frequently used it, but it played a much smaller role in his economic theory than it did in that of the neoclassicists.


139. See POLANYI, supra note 138, at 45–69.
140. See id.
142. See Veblen, Limitations, supra note 126.
143. On Clark’s marginalism, see infra Part III.D.
economic writing too often decayed into descriptive, polemic historicism.  

It never developed a rhetorically powerful and technically useable equivalent of neoclassicism’s concept of declining marginal utility, Darwin’s natural selection, or Einstein’s general theory.

At the same time, however, old institutionalism contained a set of political ideologies that was attractive to noneconomists during a period in which historians and social commentators were generally moving toward the left while neoclassical economics was moving toward the right. As a result, institutionalism has always occupied a place in American intellectual and political history that is quite out of proportion to the place it occupies in economic doctrine. By contrast, in mainstream histories of economics, the classical political economists and the prominent neoclassicists are given generous treatment, while institutionalists are ignored or barely mentioned.

The early institutionalists’ main critiques of neoclassicism, and marginalism in particular, were that it was too abstract and lacked empirical content; that it elevated rationality over evolution as a device for explaining human choice; and perhaps most importantly, that it was overly focused on market exchange as the dominant method of social interaction. The last point is crucial because it reveals a debate about the “scope” of economics that is quite different from the scope that Lionel Robbins envisioned.

Most of the institutionalists, including the legal realists, were very conscious of the power that some social institutions wield over individuals. They were typically hostile toward the neoclassical search for models of competition in which individual actors were powerless and everyone was regarded as having infinite freedom to make any exchange that he or she pleased. When institutionalists

---

145. Joseph Schumpeter’s critique was particularly harsh, notwithstanding his own sympathy with historical methods generally. He lumped the institutionalists together with practical nonacademic economists who wrote about policy and had no use for theory:

They . . . looked upon ‘marginalism’ as a sort of speculative philosophy or as a new sectarian ‘ism’ which it was precisely their business to eliminate by what they considered truly scientific and realistic research. Hence they passed, in methodological and programmatic pronouncements, all sorts of sweeping judgments upon it. On the surface, the result was bedlam, especially in Germany and in the United States—a multitude of discordant voices, all of which seemed to testify to the presence of an impasse.


147. See, e.g., BLAUG, supra note 10, at 700–03 (briefly describing American institutionalism).

148. See supra notes 85–87 and accompanying text (arguing that Robbins narrowed the scope of economics so as to exclude psychology); James Ronald Stanfield, The Scope, Method and Significance of Original Institutional Economics, 33 J. ECON. ISSUES 231, 233–36 (1999).
thought about the economic system, they tended to see monopoly as a much bigger problem than the neoclassicists did. They also tended to view business corporations as institutions that wielded great power and were inclined to use it irresponsibly—a theme that stretches across institutionalist and legal realist writings from Veblen149 to Richard T. Ely,150 Robert L. Hale,151 John Maurice Clark (and his work on social control of business),152 John R. Commons,153 and culminates in Adolf Berle and Gardiner Means’s influential New Deal book, The Modern Corporation and Private Property.154

Finally, at least one important American institutionalist, John R. Commons, was heavily concerned with the relationship between institutions and transactions—which gives him a unique place in New Institutional Economics (NIE).155 Commons’s conception of “transactions” was broader than the conception generally used in NIE. In a lengthy discussion of transactions in Legal Foundations of Capitalism, written in 1924, Commons developed a theory of institutions as “going concerns” governed by “working rules,” and whose principal resource deployment device was the “transaction.” However, for Commons, transactions came in three kinds: rationing transactions, managerial transactions, and bargaining transactions.156 Both rationing and managerial transactions were hierarchical in the sense that they were between a legally superior person and a legally inferior one. Rationing transactions were events such as a legislature’s passage of a tax or tariff bill, or a judicial decision that transferred wealth from one person to another. By contrast, managerial transactions were mainly private and referred mainly to the employment relationship, although it also included other institutions such as slavery. Only “bargaining transactions” involved voluntary contractual exchanges between legally equal persons.157

149. VELEN, BUSINESS ENTERPRISE, supra note 126; VEBLEN, ABSENTEE OWNERSHIP, supra note 126.
151. Robert L. Hale, Coercion and Distribution in a Supposedly Non-Coercive State, 38 POL. SCI. Q. 470 (1923); see also Frank, supra note 141, at 481–89 (describing evolution of coercion in historical cultures, including legal rules among other structural norms and taboos).
152. John M. Clark, Social Control of Business (1926) [hereinafter CLARK, SOCIAL CONTROL].
153. John R. Commons, Legal Foundations of Capitalism (1924) [hereinafter Commons, LEGAL FOUNDATIONS].
154. BERLE & MEANS, supra note 46.
155. The writings of Oliver E. Williamson are representative of NIE. See, e.g., Williamson, supra note 9, at 12, 26, 45, 220, 251, 371; Oliver E. Williamson, The Economic Institutions of Capitalism 3–10 (1985); Williamson, supra note 42, at 3, 6, 24, 254.
156. Commons, Institutional Economics, supra note 144, at 59–72.
157. Id. For some of those sympathetic with institutionalism, such as Richard T. Ely and Robert L. Hale, even bargaining transactions were seen as coercive because they tended to reinforce pre-existing distributions of power and increased the maldistribution of wealth in the process. See, e.g., Ely, supra note 150; Hale, supra note 151. These views eventually
Commons defined institutions as organizations that manifest collective action “in control of individual action.”158 In order to exercise such control, institutions had to develop “working rules,” which are rules that specify transactions, or movements of rights or resources among individuals, that create order within the institution. Once again, Commons’s definitions were very broad. Common law contract rules are certainly working rules, but so are language, customs, and habits. These “rules are necessary and their survival in history is contingent on their fitness to hold together in a continuing concern the overweening and unlimited selfishness of individuals pressed on by scarcity of resources.”159

In 1925, Commons wrote an essay on “law and economics” which appeared in the *Yale Law Journal.*160 In his essay he urged greater study of the relationship between market transacting and the legal system. In particular, he observed, the coalescence of these two institutions, the economic market and the legal system, created the possibility of transactions in absenta by substituting for commodities themselves a concept of legal control over them. In “primitive markets,” commodities and consideration exchanged hands simultaneously, but in modern economies the legal system facilitated mechanisms by which the transaction and the movement of the resource itself could be widely separated events.161

One of institutionalism’s most durable contributions was that it increased economic analysis of the legal system—particularly the role of property rights in economic exchange. Almost simultaneously, legal scholars at elite institutions such as Yale and Columbia began to pay more attention to economic markets and to the relation between markets and legal rules (both common law and statutory).162 To be

---

158. See *Commons, Institutional Economics,* supra note 144, at 72 (“From this universal principle of collective action in control of individual action by different kinds of sanctions arise the ethical and legal relations of rights, duties, no-rights, no-duties, and the economic relations not only of Security, Conformity, Liberty, and Exposure, but also of Assets and Liabilities.”).


162. An excellent but typically overlooked example is Lawrence Kelso Frank, whose 1924 *Columbia Law Review* article, *An Institutionalist Analysis of Law,* described the evolution of property and contract as social norms, much like tribal taboos, but which had become much more social (rather than individual) as a result of the revolution in “machine production,” which Frank believed gave rise to legal rules that emphasized cooperation more than individualism. See Frank, supra note 141. Frank argued that the movement for the new
sure, they were by and large not marginalists. Further, they tended to see various institutions, such as the legal system and the market system, as coordinate and as having their own values and operating rules. That is, they did not view all institutions through the lens of price theory.163

Legal realism was the companion of institutionalism, but only in legal theory. While legal realists championed the study of law and economics, they rarely promoted neoclassical economists. They were by and large not marginalists, and in any event they did not accept marginalism as an exclusive or even dominant theory of value. Their theory of choice was heavily behaviorist, and they were Darwinian in their thinking about the relationship between people and their environment. When they looked at traditional markets, they tended to see a great deal of monopoly, market failure, and coercion rather than competition. The legal realists tended to view markets and governments as alternative ways of allocating resources, and they tended to have at least as much confidence in government as in markets as resource allocation devices. Finally, they preceded, rejected, or ignored the severe constraints on neoclassical welfare economics posed by the ordinalist revolution. As a result, their approach to welfare rested mainly on Darwinian objective welfare judgments. However, they facilitated the interdisciplinary study of law and economics more than any group of intellectuals that preceded them.164

III. COASE: FROM NEOCLASSICISM TO NEW INSTITUTIONALISM

A. Coase and Old Institutionalism

Throughout his career Coase professed a firm belief that institutions do matter, and that neoclassical economics tended toward too much abstraction. He shared these two fundamental principles with the first generation of institutionalists. At the same time, however, Coase rejected the institutionalists’ severe limitations on marginalist analysis. Beginning with The Nature of the Firm in 1937,165 Coase used exclusively marginalist criteria to define the behavior of economic agents and continued to do so throughout his career, including his analysis of bargaining and the common law in The Problem of Social Cost.166 For Coase, as for any neoclassicist, the business firm was no different than the individual equating utilities at the margin, except that the firm’s utilities were in fact profits and losses that could be quantified by price theory. In the hands of later followers, Coasean institutionalism reached other types of institutions as well, including governments,

Restatements of the law, which were just underway at the time of his writing, were an effort to bridge the hyper-individualistic tendencies of nineteenth-century law with the realities of coordinated production in the machine age. He saw this as a failed attempt to “pour new wine into old bottles.” Id. at 491–92. He ultimately concluded that legislative rather than common law rules were the preferred route for legal change. Id. at 495.

164. See Hovenkamp, Knowledge About Welfare, supra note 136.
interest groups and lobbyists, families, religion, labor unions, educational institutions, and clubs, to name a few.\textsuperscript{167}

Second, to the extent he expressed his views on the subject at all, he sided with ordinalism in the debate over interpersonal utility comparisons. This separated him from the first generation of institutionalists, whose theories of welfare were generally not based on marginal utility functions at all. While Coase’s writing was dominated by concerns for identifying welfare-enhancing outcomes, welfare in his conception concerns two things. First, it refers to the individual rankings of an economic actor’s own internal priorities, as in \textit{The Nature of the Firm}. Second, it refers to the results of unconstrained private bargaining, as in \textit{The Problem of Social Cost}. Within this system, welfare gains occur only when parties are able to bargain at suitably low cost, or else when the legal system emulates the outcome that people would have achieved had they been able to bargain. There is not much place for any notion that involuntary transfers from, say, the wealthy to the poor will increase welfare. Coase was consistently opposed to government intervention in the economy. In sum, Coasean analysis is purely neoclassical in the sense that it identifies the welfare maximizing outcome as the one that people would choose through voluntary exchange, provided that transaction costs are suitably low.

While predecessors such as Pigou wrote extensively about the costs of moving resources, as a pre-ordinalist Pigou tended to look at the full range of interventions by which resources are moved.\textsuperscript{168} The importance of the Coasean shift in emphasis is difficult to exaggerate. Although Coase had little to say about the debate over interpersonal utility comparisons that was swirling around him at the London School of Economics,\textsuperscript{169} his reduction of the cost of moving resources to “transaction costs” meant that Coasean economics measured welfare entirely by reference to individual decisions that were efficient in the Pareto sense. This amounted to nothing short of a “neoclassical takeover” of institutionalism.

\textbf{B. The Business Firm and Institutionalism, Old and New}

For most institutionalists, the business firm deserved special study because it was such an obvious exception to the neoclassical supposition that resources are allocated through markets. Firms were in fact aggregations of individuals, collections of plants and other productive units related both horizontally and vertically, and composed of owners, managers, and workers. At least in private commerce, most resource redeployment occurred with the involvement of at least

\begin{itemize}
  \item \textsuperscript{168} See supra notes 50–51 and accompanying text; see also Hovenkamp, \textit{Coasean Markets}, supra note 18.
  \item \textsuperscript{169} See supra notes 43–70 and accompanying text.
\end{itemize}
one firm. Nevertheless, the relationships among the numerous individuals that functioned inside a firm hardly fit the neoclassical idea of a market.

What provoked so many institutionalists was that in their view the neoclassicists solved this problem by murder. Neoclassical economics largely ignored questions about the nonmarket relationships and decision making that go on inside the firm. To them, the firm was simply an “economic agent”—a black box that transacted in the market in the same way that a natural person transacted, but whose internal preference orderings were no more scrutable than the unevaluated but sacrosanct preference orderings of a single individual. The one big difference between the firm and the human being was that the firm had a profit function that maximized wealth, while the human being maximized utility. This explains why the ordinalist revolution never affected price theory and industrial organization in the way it affected welfare economics and the economics of public policy. A business firm’s wealth could be specified in cardinal units, such as pounds or dollars, and could be compared from one actor to another.

The brilliance of Coase’s *The Nature of the Firm* was to merge marginalism and institutionalism by applying marginal analysis to the firm’s internal decision making and thus to questions about its size and structure. However, Coase was not the first economist to conceive of the problem in this fashion. By 1937, there was already an important marginalist literature on the internal workings of the firm and the determinants of its size, including the rationales for vertical integration, or a firm’s decision to produce an input for itself rather than purchase it on the market. One important older linkage between marginalism and firm structure was Alfred Marshall’s 1919 book *Industry and Trade*, described above.\(^\text{170}\) Much of the literature on vertical integration related to production cost savings.\(^\text{171}\) There was also a strong tendency, stretching from the 1910s through the New Deal, to see vertical integration as a monopoly problem.\(^\text{172}\)

The ideas of several economists who wrote in the 1920s deserve special mention, however, for they expressly related questions about firm size and structure to the issues of incentives and, in some cases, the costs of using the market. These were Lawrence Frank’s discussion of coordination problems attending the rise of machine production, John Maurice Clark’s pathbreaking study of *Overhead Costs*, and Pigou’s, Kaldor’s, and E.A.G. Robinson’s important marginalist analyses of the equilibrium business firm.

---

170. MARSHALL, INDUSTRY AND TRADE, supra note 38; see supra note 39 and accompanying text.


C. Lawrence Kelso Frank

Lawrence Kelso Frank’s theory of vertical integration, developed in an essay written in 1925, was based on a comparison of the costs of internal production as opposed to market acquisition. For Frank, an important cost of using the market in the machine age was the cost of coordination.\(^{173}\) He argued that firms had been vertically integrated prior to the industrial revolution because hand craftsmen tended to perform all of their operations internally, from obtaining source materials to producing and marketing a product.\(^{174}\) The rise of machine production, however, led to much larger output but much smaller vertical size, as machines tended to perform discrete, repetitive functions at high speed.\(^{175}\) Machines also had to be kept in production at rates that were not driven by immediate demand, however, and this led to a market coordination problem. The result, Frank explained, was two-fold: development of more elaborate contractual coordination and the rise of modern vertical ownership integration. Ultimately, he argued, vertical ownership integration proved superior in most cases.\(^{176}\) But Frank drew these conclusions, which clearly anticipated Coase, comparing the costs of the market with those of vertical ownership:

> It is instructive to contrast this method of co-ordinating the several stages of production in an industry, with reliance upon the pecuniary relationships of buying and selling, where production is governed by fluctuating market prices, a multitude of individual contracts or orders subject to cancellation, the guesses and expectations of salesmen, and above all by the widespread practice of speculative purchasing which alternately overbuys and then refrains from buying while using up accumulated stocks.\(^{177}\)

D. Clark and Overhead Costs

John Maurice Clark was one of the most incisive economists of the business firm during the interwar period. Clark sat somewhat uncomfortably between the marginal utility theory, of which his father John Bates Clark had been a champion, and the institutionalism of the 1920s, of which he considered himself a member.

\(^{173}\) Lawrence K. Frank, *The Significance of Industrial Integration*, 33 J. Pol. Econ. 179 (1925).

\(^{174}\) *Id.* at 179–80.

\(^{175}\) *Id.* at 180 (“[T]he machine process split up production into an ever growing number of separate processes, separate because of the invention of new techniques and new machines for performing each step in the formerly unified handicraft operations.”).

\(^{176}\) *Id.* at 185–86 (“For undoubtedly vertical integration is an attempt to bring together under one management the separate stages of the industrial process which technically require unified direction and control. Since this technical requirement cannot effectively nor continuously be met through buying and selling of goods between separately owned stages, however ingeniously and elaborately those pecuniary operations be conducted, it has become both feasible and desirable to bring a number of consecutive stages of production under one managerial control.”).

\(^{177}\) *Id.* at 190.
Like many institutionalists, he was a pragmatist, more interested in workable solutions than in doctrinal purity. In 1921 he criticized neoclassical economics for being too “Euclidean,” and for using its mathematics to obfuscate rather than illuminate social issues.178 His 1926 book entitled Social Control of Business was heavily influenced by the social science movement and intended to show the need for more social concern in and public regulation of private enterprise.179 In 1930 he argued that neoclassical economics had become far too individualistic and needed to articulate a social agenda.180 Later in his career, he wrote a formative policy essay on “workable competition,” designed to create a set of antitrust rules that would avoid being continually frustrated by an overtly pure conception of perfect competition.181

In the early 1920s, Clark made a sustained attempt to use marginalist analysis to explore the internal decision making of the business firm. The book, entitled Studies in the Economics of Overhead Costs, dealt with a problem that had vexed neoclassical economics since the time of Marshall, which was an antinomy between the theory of perfect competition and the presence of fixed costs.182 Prices are always determined by decision making on the margin, and under perfect competition they are driven to marginal cost. That is, a firm will produce one more unit if the incremental cost of making and selling it is no greater than the market price. Beginning with the “barren tautology” that the competitive price equals marginal cost, Alfred Marshall chose to ignore most of the internal workings of the firm and instead introduced the hypothetical idea of a “representative firm” that held in common the unexplored but efficient characteristics of firms operating in a particular market.183 This permitted Marshall to talk about firms without troubling himself very much about the differences among them.

Under the Marshallian model, however, there was no profitable equilibrium for firms in competitively structured markets with significant fixed costs. A fixed cost is an investment in something like land, plants, equipment, or research and development that does not vary with production and that must be paid whether or not the firm produces anything. Fixed costs create the concept of “overhead,” or costs unrelated to the incremental costs of making an additional unit and selling it. Within Marshall’s model, competition was thought to become “ruinous” as each firm cut its price to marginal cost without having enough left over to pay off fixed costs. Firms would either go out of business until only a single monopoly firm remained or else they would be forced to collude. This “fixed cost controversy” had

179. CLARK, SOCIAL CONTROL, supra note 152; see Herbert Hovenkamp, United States Competition Policy in Crisis: 1890–1955, 94 MINN. L. REV. 311, 323–24 (2009) [hereinafter Hovenkamp, Competition Policy].
180. Clark, Socializing, supra note 124.
181. J.M. Clark, Toward a Concept of Workable Competition, 30 AM. ECON. REV. 241, 241–43 (1940); see Hovenkamp, Competition Policy, supra note 179, at 323, 342–43.
182. J. MAURICE CLARK, STUDIES IN THE ECONOMICS OF OVERHEAD COSTS (1923) [hereinafter CLARK, OVERHEAD COSTS].
183. Id. at 13.
a major impact on antitrust policy in the United States during the early decades of the twentieth century.\textsuperscript{184}

Clark observed that fixed costs were an increasingly important aspect of the business firm under modern methods of mechanized production. \emph{Overhead Costs} explored the many ways that firms make choices so as to maximize their position given the presence of fixed costs, and some of these choices pertained to vertical integration, also the subject of Coase’s \emph{The Nature of the Firm}.

Clark’s book claimed a great deal. Market demand was constantly in flux, but in industries with significant fixed costs productive capacity was inelastic, or not subject to much variation.\textsuperscript{185} Business production could not expand or contract on a moment’s notice in response to market conditions. Once machinery was in place, high output led to low costs but also to a potentially disastrous loss of investment in the case of a shutdown. Examination of these problems provided Clark, just as Lawrence Frank, a basis in fixed costs for a theory of business cycles, which explained the tendency of capitalist economies to swing from boom to bust.\textsuperscript{186} Clark applied his theory first to industries with obviously high fixed costs, such as railroads\textsuperscript{187} and public utilities,\textsuperscript{188} but then generalized to all of modern manufacturing. He found that the problem of fixed costs affected numerous decisions concerning capital, labor, and selection of raw materials. “From being a mere exception to the general laws of value and efficiency,” Clark wrote, the study

\begin{itemize}
  \item \textsuperscript{185} Clark, \textit{Overhead Costs}, supra note 182, at ix (introducing the book as “a study of discrepancies between an ever fluctuating demand and a relatively inelastic fund of productive capacity, resulting in wastes of partial idleness, and many other economic disturbances. Unused capacity is its central theme”).
  \item \textsuperscript{186} Frank made similar observations in \textit{Institutionalist Analysis}. Frank, supra note 141, at 491. He noted that, historically, under hand-production there was a close relationship between output and prices because production was responsive to demand virtually on a unit-by-unit basis. By contrast, machine production led for the need of machines that kept running at high output regardless of demand. This fact, plus the offsetting correctives made by managers, led to booms and busts as industries swung between over- and under-production. Joseph Schumpeter’s massive study of business cycles came almost two decades later but developed many of the same points with copious historical examples. Joseph A. Schumpeter, \textit{Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process} (1939). While Schumpeter poured his soul into the book, it is largely regarded today as his least successful. See Thomas K. McCraw, \textit{Schumpeter’s Business Cycles as Business History}, 80 BUS. HIS. REV. 231, 231 (2006).
  \item \textsuperscript{187} Clark, \textit{Overhead Costs}, supra note 182, at 258–97.
  \item \textsuperscript{188} \textit{Id.} at 318–34.
\end{itemize}
of overhead costs “has grown to be a large and important section of economic principles.”

Clark also discussed the relationship between overhead costs and the nature of the firm. Many of his observations about why firms integrate vertically would resemble those made by Coase fifteen years later in *The Nature of the Firm*. Some of them anticipated the transaction costs economics of vertical integration and contracting in the 1970s. One consequence of vertical integration, Clark observed, is that it often changed the nature of a firm’s costs from variable to fixed; by contrast, the market tends to preserve costs as variable and as a result fails to recognize the true impact of overhead costs. A consistent theme in *Overhead Costs* is that fixed costs and high-volume machine output means that manufacturing and orders cannot be synchronized on a day-by-day basis. Machines must either run or be shut down. While in the long run there might be equilibrium between supply and demand, the modern firm could expect to face periods of over- and under-supply. The greater the proportion of a firm’s fixed costs, the more extreme these cycles would be. This need for constancy of output in the face of cycling business conditions created a tremendous coordination problem, both upstream and downstream. Clark conducted a simple thought experiment, imagining a world in which all production and distribution of a finished good was in the hands of a firm owned by its workers. In that setting, he concluded, many of these coordination problems would go away because this hypothetical firm would be in a better position to shift productive resources to the place where they would be needed the most.

But in the real world coordination problems are substantial. Clark argued that they provided the basis for a manager’s decision about when to contract out and when to produce internally. Clark anticipated a theory of information “impactedness,” namely, that each manager knows his own business far better than the business of others and that learning about the others is costly. As a result, a business firm might find that internal production would be more reliable than trusting another firm as seller or purchaser of an intermediate product. Not only would this address the coordination problem, Clark suggested, but it would also reduce collateral costs of reliance on the market, namely, the “work of negotiation, bargaining, higgling, stimulating demand (on the part of the seller)[.]

---

189. *Id.* at ix.
191. *Clark, Overhead Costs,* supra note 182, at 23–24, 403–07.
192. *Id.* at 386.
193. *Id.* at 402.
196. *Id.* at 137 (“Another gain from integration arises, in the shape of great reliability in the supplying of materials. The two concerns adapt their process to each other, and the supply of materials, both in quality and regularity, can be more carefully suited to the needs of the user . . . .”).
testing qualities (on the part of the buyer), and much of the other work of buying and selling. Clark also argued for a concept of “intellectual overhead,” which recognized that information is not only impacted, but it is also fixed in that, once obtained, it can be stretched over any volume of output. This made having one’s own information more valuable as output grew larger.

Business cycles entailed that market prices would not necessarily move in relation to demand. This affected the firm not only as a seller of its own product, but also as a buyer of raw materials. Fixed costs could also lead to inefficient risk distribution, Clark argued, because those firms who had mainly variable costs would be in a position to shift the risks of over- and under-supply to firms with high fixed cost. This created a need for longer-term contracting and special attention to the problem of “renewals,” which must be planned in such a way that firms would not be able to take advantage of one another. Clark also noted that the degree of specialization of high volume machinery exacerbated coordination problems and the ability of firms to behave strategically in buying and selling. These would be lessened to the extent that machines are adaptable to changing production. “In some cases, however, a change in the product calls for large investments in new equipment.” Nevertheless, the “regularizing of renewals and extensions” is “an ever-present need.” Finally, vertical integration into consumer sales would likely produce dealers that have a “much more complete harmony of interest with the producer,” thus making them more effective resellers.

Clark also observed that “[n]ot all the forces are working in the direction of integration.” Economies of scale may incline a firm not to integrate into a market where its own needs would be too small to yield efficient production. Also cutting against vertical integration was specialization held by other firms, which would operate in areas where trained producers could do something better than a firm could do it internally.

The opening sentence in an important article by Klein, Crawford, and Alchian in the 1970s acknowledges “Coase’s fundamental insight that transaction, coordination, and contracting costs must be considered explicitly in explaining the extent of vertical integration.” The authors do not mention Clark, but he had, in fact, developed a comprehensive theory fifteen years earlier that contracting and coordination costs determine when a firm will choose outside procurement or internal production. These choices related entirely to the firm’s search to maximize

197. Id. at 137.
198. Id. at 119–22.
199. Id. at 399–400.
200. Id. at 138.
201. Id. at 137–38.
202. Id. at 138.
203. Id.
204. Id. at 138–39 (citing Lewis H. Haney, Integration in Marketing, 10 AM. ECON. REV. 528 (1920)).
205. Id. at 140.
206. Id. at 139.
207. Id. at 140–41.
208. Klein et al., supra note 190, at 297.
its own value within an environment that included fixed costs and specialization, as well as shortness of information about the positions of bargaining partners. Clark’s analysis was not tied into as neat a package as that of Coase, but the fundamental theory was the same.

Clark’s *Overhead Costs*, just as much as Coase’s *Nature of the Firm*, broke the business firm apart and explored its inner decision making. For both, the premise was that managers set out to maximize the firm’s value. In that sense Clark’s *Overhead Costs* was firmly marginalist just as it was firmly institutionalist. While Coase’s *Nature of the Firm* is given considerable credit for producing non-monopolistic explanations for business decisions, Clark’s *Overhead Costs* did the same thing—pointing out, for example, that pricing above short run cost and price discrimination are not unique attributes of monopoly, but can exist in any market with fixed costs.\(^{209}\) The pervasive theme of *Overhead Costs* is that many of the pricing anomalies that represented deviations from perfect competition in fact had completely nonmonopolistic explanations when durable fixed-cost investment was relatively high.

As Harvard’s T.H. Sanders wrote in a review, Clark’s discussion addressed the problem of those “little boxes” that had characterized neoclassical discussions of microeconomics.\(^{210}\) More particularly, Clark raised the question whether “the experiences of the individual firm [are] of interest to the economist; or is he concerned only with the generalizations which can reliably be made from the observation of many individuals firms?”\(^{211}\)

**E. Pigou, Kaldor, and Robinson: the Firm in Equilibrium**

Pigou, who later became Coase’s frequent target, wrote an important article on the “equilibrium firm” in the late 1920s.\(^{212}\) Pigou’s focus, unlike Coase’s, was mainly on horizontal size rather than vertical integration. Pigou showed that when the supply price of the industry is greater than the marginal cost of the individual firm, that firm will tend to expand until the two are equalized. When the industry’s supply price is lower than the firm’s costs the firm will contract. As a result, the only firm that is in equilibrium (assuming that the market as a whole is in equilibrium) is the one whose marginal cost equals the industry supply price.\(^{213}\) Pigou’s article related expansion directly to cost, although it did not distinguish transaction costs from other kinds of costs. In an important article on firm equilibrium, Nicholas Kaldor added to this that the size of firms is limited to the

\(^{209}\) *Clark, Overhead Costs*, supra note 182, at 416.


\(^{211}\) *Id.*


\(^{213}\) Pigou elaborated on this work in *The Economics of Welfare*, *Pigou*, supra note 4, pt. II, ch. 11, §§ 3–12.
extent that production costs rise.\footnote{Nicholas Kaldor, \textit{The Equilibrium of the Firm}, 44 \textit{ECON. J.} 60 (1934).} Once the firm’s costs equal those of the market, the firm will no longer grow.\footnote{\textit{Id.}}

In 1931, Cambridge economist E.A.G. Robinson published his seminal book on industrial organization, \textit{The Structure of Competitive Industry}. Robinson’s book looked beyond the individual plant to problems of management.\footnote{E.A.G. Robinson, \textit{The Structure of Competitive Industry} (1931).} When Pigou and Kaldor spoke of costs they generally assumed a firm that operated a single plant. Just as production, management was also subject to both economies and diseconomies of scale.\footnote{Id. at 115 (citing United States Steel as an example of a firm that gained economies by operating multiple plants scattered across the country).} Management costs differed from plant-specific costs, however, in that they applied to the firm as a whole.\footnote{Robinson elaborated on this point in Austin Robinson, \textit{The Problem of Management and the Size of the Firm}, 44 \textit{ECON. J.} 242 (1934).} To the extent that management costs per unit rose as firms grew in size, these costs placed a limit even on the size of multi-plant firms. Only by taking management as well as firm-specific production costs into account could one come up with a limiting measure on firm size.\footnote{Coase, \textit{Nature of the Firm}, supra note 26, at 386.}

The work done by Pigou, Kaldor, and Robinson were exercises in pure marginalism, not empirical studies of firms. With increasing degrees of complexity, they related the size of a firm to its increasing internal costs of production and management, as opposed to the supply costs of the market as a whole, which were presumed to be constant.\footnote{See Sraffa, supra note 212, at 535–36.} For example, if market supply costs were a constant $10 per unit, a firm whose current costs were $8 per unit would grow until its own costs rose to $10, but not further. Alternatively, a firm whose costs exceeded $10 would contract until its costs fell. A firm would be in equilibrium when its own costs precisely equaled the costs of the market as a whole.

Given these contributions, Coase certainly exaggerated in the opening paragraphs of \textit{The Nature of the Firm}, when he suggested that his paper was the first to define the scope of the firm by using a theory of marginal substitution.\footnote{See Joan Robinson, supra note 89, at 69.} The work done by Pigou, Kaldor, and Robinson all related the size of a firm to its marginal costs as opposed to those of the market as a whole, and found equilibrium at the point where firm costs and market costs were in equipoise. What Coase added was a critical tool for determining how firms grow with a higher degree of specificity. He simply observed that markets themselves are costly, and firms are always seeking to minimize costs. As long as the marginal costs of internal production are lower than the marginal costs of using the market, the firm will produce internally, and vice-versa. Coase’s analysis applied to every decision a firm made, from relatively “macro” decisions about whether or not to build a new

---

\footnote{Nicholas Kaldor, \textit{The Equilibrium of the Firm}, 44 \textit{ECON. J.} 60 (1934).} \footnote{\textit{Id.} Kaldor also concluded that increasing firm costs were inconsistent with perfect competition. \textit{Id.} at 72–73.} \footnote{E.A.G. Robinson, \textit{The Structure of Competitive Industry} (1931). E.A.G. Robinson, also a Professor of Economics at Cambridge, was married to fellow professor Joan Robinson. See Joan Robinson, supra note 89, at 69.} \footnote{\textit{Id.} at 40–42.} \footnote{\textit{Id.} at 115 (citing United States Steel as an example of a firm that gained economies by operating multiple plants scattered across the country).} \footnote{Coase, \textit{Nature of the Firm}, supra note 26, at 386.}
plant to extremely “micro” decisions about whether to sweep one’s own floors rather than hiring a cleaning service, or whether to hire an employee lawyer or engage outside counsel.\footnote{See Hovenkamp, Coasean Markets, supra note 18.}

At the same time, for all of Coase’s subsequent rhetoric about actual observation and the real world,\footnote{See, e.g., Coase, Nobel Prize Lecture, supra note 8 (calling for more empirical study of the structure of the firm, and lamenting “blackboard economics” that reduces the firm to a “black box”); see also Coase, New Institutional Economics, supra note 8, at 72 (“Mainstream economics, as one sees it in the journals and the textbooks and in the courses taught in economics departments has become more and more abstract over time, and although it purports otherwise, it is in fact little concerned with what happens in the real world.”).} very little in \textit{The Nature of the Firm} concerns empirical observations about anything. In fact, Coase’s mode of analysis was the same as that of his neoclassical predecessors. He assumed that the firm seeks to maximize its value and that both internal production and markets impose costs, but that these costs are not necessarily the same. A firm directs its activities between internal production and external procurement until, at the margin, the value of one equals the value of the other. Coase never “verified” this conclusion and did not need to: as was the case for Clark’s work on fixed costs, Pigou’s theory of the equilibrium firm, or Robinson’s theory of management costs, Coase’s conclusions flowed precisely from marginalism plus the neoclassical theory that the firm is a maximizer of its own value. As Coase concluded:

\begin{quote}
To determine the size of the firm, we have to consider the marketing costs (that is, the costs of using the price mechanism), and the costs of organizing of different entrepreneurs and then we can determine how many products will be produced by each firm and how much of each it will produce.\footnote{Coase, \textit{Nature of the Firm}, supra note 26, at 403.}
\end{quote}

Coase’s only attempt to provide some real world verisimilitude to this powerful theory was a concluding discussion of the law of master and servant. Relying on English lawyer Francis Raleigh Batt’s treatise on employment law, Coase observed that the common law distinguished between market relationships and agency, or command and control relationships such as those between employer or manager and employee.\footnote{Id. at 403–04 (relying on Francis Raleigh Batt, \textit{Law of Master and Servant} (4th ed. 1950) (1926)).} Coase later wrote at some length about the composition of \textit{The Nature of the Firm}, even suggesting that it was a kind of empirical exercise for which he initially collected data during visits to various manufacturing plants on a trip to the United States.\footnote{See Coase, \textit{Nature of the Firm: Origin}, supra note 5, at 8–9 (explaining that during his visit to the United States he took very few classes and instead “most of my time was spent in visiting businesses and industrial plants”); see also id. at 13 (noting a visit to Duquesne Power and Light Co. and the viewing of some cost curves that they had constructed); \textit{id.} at 13 (visit to General Motors and discussion of Fisher Body works).} That visit may well have inspired Coase to write what
he did, but his famous essay contains very little in the way of verification or even casual observation of business firm decision making.

Lest this be regarded as critical of Coase’s influential article, it bears mention that Coase was absolutely clear from the beginning about what he was doing. He wanted to produce a tractable definition of the firm that could be derived 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.” Whatever methodology was contained in The Nature of the Firm, it was clearly not the kind of empirical and historical investigation envisioned by the first generation of institutionalists.

CONCLUSION: FROM COASEAN INSTITUTIONALISM TO LAW AND ECONOMICS

Within neoclassical economics prior to Pigou, resources were generally assumed to be unambiguously assigned and costlessly traded. In that regime the need for a legal system was limited to the enforcement of contracts, prohibition of fraud, and control of monopoly. Except for the concern about monopoly, the neoclassicists did not trouble themselves all that much about the interaction of the legal system and the market. The rise of institutionalist economics in the early twentieth century changed that, as the early institutionalists examined the many forms of nonmarket organization through which resources move.

Today, the Coasean vision of the relationship between institutions and law has very largely swept the field. Within economics itself, New Institutional Economics (NIE) accounts for an important portion of the work that is done; it is largely considered to be in the mainstream, as evidenced by the Nobel Prize given in 2009 to Oliver E. Williamson, its best-known advocate. NIE shares in common with old institutionalism the view that organizations and institutions differ from neoclassical markets and thus are worth studying. Within this paradigm, “organizations” are associational bodies of people that have a set of management rules that control their decisions. Good examples are corporations or other business firms, governments, and clubs. “Institutions” are the sets of rules that define how decisions are made within any organization. These rules can be both formal and informal; they can have the force of law or simply of custom, habit, or mutual recognition. The common law, price theory, parental controls over children, corporate voting structure, or the cloture rules of the United States Senate are all examples.
This creation of a new institutionalism within neoclassicism has done two things at once. First, it has enabled institutionalism as a theory to have more coherence and explanatory power than the old institutionalism had. Indeed, one of the most devastating critiques of the old institutionalism was that it quickly turned into little more than historical description, offering very little analytic power.233

But the other thing that NIE has done is much more constraining. Economics is fundamentally about producing value through the creation and movement of resources. Under the constraints of marginalism and ordinalism, however, only a small subset of resource movements can be defended as socially beneficial. More fundamentally, only a small subset of such movements is thought to be within the domain of scientific economics. First are individual preference orderings and voluntary exchange. A voluntary exchange necessarily increases value to both parties (or at least, does not diminish the value to either). If an exchange affects no one else, it meets even the strict Paretian efficiency requirements. Second is price theory, which unlike welfare economics is concerned with the maximization of value in monetary terms. Third is the structure and decision making of institutions, including but not limited to the business firm. Whether regarded as an individual economic actor or as a collection, the firm’s goal is to maximize its own value. One thing that makes modern law and economics both insightful and controversial is that this same goal, the maximization of value, is used to explain nonbusiness institutions as well.

The principal effect of all of this has been to restrict the economic study of legal institutions to the role of transaction costs; “transaction” is defined to include both resource transfers that result from bargains between legally independent people and nonmarket transfers that occur within the economic agent itself. The idea that the cost of moving resources from one point to another is important to the assessment of an economy hardly originated with Coase. It was very well developed in Pigou’s work before Coase published anything.234 But writing as he did, prior to the ordinalist revolution, Pigou had a much less “contract focused” theory about welfare-increasing movements of resources. Pigou thus spoke more broadly than Coase later did about state policies that redeployed resources from lower to higher value uses.

By the time that The Problem of Social Cost was written the ordinalist victory in neoclassical economics was complete.235 Coase’s 1960 article was concerned almost entirely with the use of private transactions as resource movement devices.236 Coase simply assumed that the domain of welfare-enhancing resource movements was no broader than the domain of voluntary transactions. As a result, the “resource movement” issue largely transformed itself into a “transaction cost” issue.

It would be incorrect to describe the modern law and economics movement as a merger of neoclassical economics with the economic institutionalism of the 1920s. While the dominant branch of law and economics is neoclassical, it draws very little from the first generation of institutionalists. Nevertheless, modern law and
economics grew out of the premises that institutions count and that the legal system is an essential institution for seeing to it that resources end up in their most valuable use. Within law and economics, the initial assignment of property rights and the transaction costs of bargaining are responsible for the creation of all economic agents larger than the single individual. Every economic event becomes a study in transaction costs. Transactions costs explain why the person who prefers the new Toyota to her Chevrolet might nevertheless choose to stay with the car she already has. They explain how large the business firm, the family, or some other institution will become and how many activities it pursues internally rather than on a market. In its most formal mode, transaction costs explain why people marry rather than purchase sexual or domestic service, why they have children rather than hire outside help, why sports teams are organized into conferences and how large the conferences are, and why the common law is more efficient than legislation.

Marginalism and ordinalism were the ingredients that served to make the Coasean brand of institutionalism, as well as its law and economics, very different from the institutionalism of the pre-War years. They also explain why law and economics to this day is defended by insiders and attacked by critics for taking strong, free market, antistatist positions that are highly suspicious of involuntary redistributions of wealth. This in turn may explain why more expansive variations, such as so-called “behavioral” law and economics, have felt constrained to reject or seriously modify marginalism as a fundamental principle of human motivation.237 When human motivation and action is understood in a broader and less formal way, the rationale for institutions necessarily does the same.

---