“Exceptional and Unimportant”?

The Rise, Fall, and Rebirth of Externalities in Economic Analysis

Steven G. Medema*

Version 2.1
September 2017

* University Distinguished Professor of Economics, University of Colorado Denver. Email: steven.medema@ucdenver.edu. The author has benefitted greatly from comments on earlier drafts of this paper provided by Roger Backhouse, Nathalie Berta, Ross Emmett, Kenji Fujii, Kerry Krutilla, Alain Marciano, Norikazu Takami and participants in the CHOPE workshop at Duke University. The financial support provided by the National Endowment for the Humanities, the Earhart Foundation, and the Institute for New Economic Thinking is gratefully acknowledged.
“Exceptional and Unimportant”?  

The Rise, Fall, and Rebirth of Externalities in Economic Analysis

I. Introduction

Economists typically locate the origins of the theory of externalities in A.C. Pigou’s *The Economics of Welfare* (1920; 1932), where Pigou suggested that activities which generate uncompensated benefits or costs—e.g., pollution, lighthouses, scientific research—represent instances of market failure requiring government corrective action. According to this history, Pigou’s effort gave rise to an unbroken Pigovian tradition in externality theory that continues to exert a substantial presence in the literature to this day, even with the stiff criticisms of it laid down by Ronald Coase (1960) and others beginning in the 1960s.¹

This paper challenges that view. It demonstrates that, almost immediately after the publication of *The Economics of Welfare*, economists largely stopped writing about externalities. On the rare occasions when externalities were mentioned, it was in the context of whether a competitive equilibrium could produce an efficient allocation of resources and to note that externalities were an impediment to the attainment of the optimum. When economists once again began to take up the subject of externalities in a serious way, the very real externality phenomena—pollution, etc.—that had concerned Pigou were not in evidence. Instead, the analysis was targeted at identifying how and why externalities violated the necessary conditions for an optimal allocation of resources in a competitive system. In short, externalities were conceived very differently in the welfare theory of the 1950s than they had been in Pigou’s treatise. It was only when economists began to turn their attention to environmental and urban

---

¹ This perception can be found as early as Wellisz (1964).
problems that we see a return to a conception of externalities as real, policy-relevant phenomena. Even then, however, the approach to externality policy was anything but straightforwardly Pigovian in nature. The history of externality theory is therefore not a history of a continuous tradition but of changing conceptions of externalities, framed by changing ideas about what economic theory is attempting to achieve.²

II. Background

The origins of the concept of externality lie not with Pigou, but in the nineteenth-century. Scholars including T.R. Malthus (1798), John Stuart Mill (1848), and Henry Sidgwick (1901) looked at the world around them and observed a wide variety of economic activities and outcomes that seemed to speak against Smith’s (1976, IV.2.9) claims regarding the efficacy of the market as a mechanism for channeling self-interested behavior to the best interests of society as a whole, and the negative spillover effects attending individual self-interested actions loomed increasingly large here.³ For these writers, ‘externalities,’ as we now know them, were real and pervasive phenomena that were evidence against the claim that a system of laissez-faire, as it was often described, would facilitate the attainment of the utilitarian maxim of the greatest good for the greatest number—whether that was measured in output-based terms or a more broad-based conception of welfare. That said, both Mill and Sidgwick, the nineteenth-century writers

---

² The history presented fills a significant gap in, and also in some ways challenges, the history of externalities elaborated by Papandreou (1994, ch. 2) in his wide-ranging discussion of the externality theory. Papandreou’s approach, though, was somewhat different, focusing on the history of how economists “have characterized externality or what they thought it was,” providing, as he put it, “a family tree of the many meanings and names associated with externality” (1994, 5).

³ Of course, Smith’s position was much more nuanced than this. See, e.g., Skinner (1996), Viner (1927), and Medema (2009, ch. 1) for discussions of this aspect of Smith’s work. It bears emphasizing that The Wealth of Nations was written against a backdrop of centuries of arguments that the market system and the pursuit of self-interest within it did not redound to the best interests of society as a whole and reflected a more favorable disposition toward the market mechanism than typically found in earlier commentaries. See, e.g., Force (2003) and Medema (2009, ch. 1).
who placed the greatest emphasis on these harmful external effects (as well as beneficial ones), were hesitant to recommend State corrective action because of concerns that the governmental cure could be worse than the market disease (Medema 2009, ch. 2).

The place of external effects in economic thinking was solidified by Cambridge economist A.C. Pigou in *The Economics of Welfare* (1920; 1932), which applied the marginalism of Alfred Marshall’s *Principles* (1890) to Sidgwick’s harmful spillover effects as one part of a much more broad analysis of social welfare issues. Pigou’s key insight here was the identification of divergences between the private and social net products associated with certain classes of activity. Under “simple competition,” he said, these divergences result from “the fact that, in some occupations, a part of the product of a unit of resources consists of something, which, instead of coming in the first instance to the person who invests the unit, comes instead, in the first instance (i.e. prior to sale if sale takes place), as a positive or negative item, to other people” (1932, 174).

Pigou identified three classes of activities exhibiting such properties: (i) situations in which those investing in durable instruments of production are not the owners of those instruments, as in the case of tenancy relationships (1932, 174-83); (ii) situations in which “one person A, in the course of rendering some service, for which payment is made, to a second person B, incidentally also render services or disservices to other persons (not producers of like services), of such a sort that payment cannot be exacted from the benefited parties or compensation enforced on behalf of the injured parties” (183), as with smoke nuisances,

---

4 Marshall, for his part, made only passing reference to external effects of this nature. See Marshall (1890/1960, 166-67). On Sidgwick’s influence on Pigou, see Aslanbeigui (1995) and O’Donnell (1979). It should also be noted that Pigou’s *Economics of Welfare* was, in the essentials, a revised version of his *Wealth and Welfare* (1912), with the latter volume reproducing most of the analysis found in the former, including that related to externalities.

5 The citations here are to the 1932 edition of Pigou’s treatise, the analysis of which, with one exception (noted below) is identical in the essentials to the 1920 edition.
lighthouses, scientific research, Veblen effects in consumption, and women’s work in the factories immediately preceding and following childbirth (1932, 183-96); and (iii) situations in which there exist increasing or decreasing returns at the industry level, as when the expansion of production in one industry has positive or negative spillovers on production in other industries (213-28). These three classes of activities were lumped together under the heading “external economies and diseconomies”—Alfred Marshall’s terminology for the third class of activities—in the decades that followed, and we will use that terminology or the shorthand, “external effects,” to describe this broader group. It is the first two of these categories, though, that overlap fairly neatly with the modern conception of “externality”—the focus of the present paper—and when the term “externality” is used here, it is in that more narrow sense.

Each of these situations, for Pigou, represented an instance in which Smith’s invisible hand could not be relied upon to maximize the “national dividend” (in essence, the value of output in society). As Pigou pointed out in his discussion of uncompensated services and disservices,

It is plain that divergences between private and social net product of the kinds we have so fare been considering cannot, like divergences due to tenancy laws, be mitigated by a modification of the contractual relation between any two contracting parties, because the divergence arises out of a service or disservice rendered to persons other than the contracting parties. (1932, 192)

---

6 Papandreou (1994, 22) suggests that it was Samuelson (1948) who made the connection of external effects to consumption—in essence, attaching “Veblen effects” to the concept to it—but Pigou had done so in 1920 and others, such as Meade (1945), Tintner (1946), and Reder (1947) had also made this connection prior to Samuelson.

7 The original elaboration of this problem of increasing and decreasing returns, though not painted in terms of divergences between private and social net products, owes to Marshall (1890/1960), and it was Marshall’s terminology of external economies and diseconomies, rather than Pigou’s divergences between marginal private and social net products, that dominated the discussion prior to 1960.

8 Pigou allowed that there were instances in which tenancy-investment-related divergences could be dealt with through adjustments in tenancy contracts.
As such, argued Pigou, there was a case to be made for State action that would remedy the divergence through the use of “extraordinary encouragements” or “extraordinary restraints,” the “most obvious forms of which,” he said, are “bounties and taxes” (192). For Pigou, these divergences between the private and social interest were not isolated occurrences but, instead, were endemic to and pervasive within the market system, and the most significant of these required government corrective action.9

Though Pigou was focused on economic welfare, as measured by the national dividend, he was convinced that, in most instances, activities which promoted economic welfare also promoted non-economic welfare—his greater concern. In other situations, however, increased economic welfare may be obtained only at the expense of non-economic welfare, with the loss on the latter front potentially outweighing the gains on the former, as in the case (for Pigou) of activities which achieved those gains at the expense of the laboring class and the poor.10 As such, the elimination of certain types of divergences between private and social net products served goals beyond the merely economic, illustrating that economics was, at times at least, a “handmaid to ethics” (Pigou 1925, 82).11

Pigou at once emphasized that those phenomena which we now classify as “externalities” represented an important and policy-relevant social problem, central to the analysis of economic

---

9 “It is as idle to expect a well-planned town to result from the independent activities of isolated speculators as it would be to expect a satisfactory picture to result if each separate square inch were painted by an independent artist. No ‘invisible hand’ can be relied on to produce a good arrangement of the whole from a combination of separate treatments of the parts. It is, therefore, necessary that an authority of wider reach should intervene and should tackle the collective problems of air and of light, as those other problems of gas and water have been tackled” (Pigou 1932, 195).

10 This point had also been emphasized by Sidgwick, who was concerned both that certain activities interfered with the maximization of the value of output and that not all activities that enhanced the value of output were welfare-increasing. See Medema (2009, 42-43), as well as O’Donnell’s (1979) analysis of the commonalities between the welfare theories of Sidgwick and Pigou.

11 See Aslanbeigui and Medema (1998) for an elaboration of Pigou’s ethical perspective and its relation to his welfare analysis, as well as the much more full discussions of Pigou’s approach in Aslanbeigui and Oakes (2015) and Kumekawa (2017).
welfare, and, by wrapping them into the framework of marginal private and social net products, provided a measure of analytical clarity for their analysis by economists. However, the broad range of phenomena that he lumped together as instances of divergence between private and social net products also had the effect of creating a conceptual muddle surrounding “external economies and diseconomies.” The efforts at clarification and untangling played out in the literature over the next several decades with only modest success, stimulating Tibor Scitovsky to remark in 1954 that “The concept of external economies is one of the most elusive in the economic literature” (1954, 143). But their elusive nature is not the only important feature of the post-Pigou literature on external effects. For, in spite of the emphasis laid onto these phenomena by Pigou, they largely disappeared from the literature until the mid-1950s. And such mention as was made of externalities was of a very different nature than we find in the writings of Pigou and his predecessors.

III. The Lacuna

We can see this lacuna in the literature at two levels. First, we find virtually no discussion of Pigou’s first two classes of divergence in the journals between 1920 and the mid-1950s.12 Such references as exist tend to be (literally or figuratively) footnotes to discussions of increasing and decreasing returns—passing references to smoke nuisances, congestion, and over-exploitation of natural resources—in articles by Frank Knight (1924), Jacob Viner (1932), Richard Kahn (1935), and Howard Ellis and William Fellner (1943). Even Knight’s article, “Some Fallacies in the Interpretation of Social Cost,” now considered a classic in the theory of externalities and part of

12 Papandreou’s (1994) discussion of the history of externality analysis, focusing on economists’ various conceptualizations of externalities during this period, does not pick up on the important point that externalities had only a minimal presence in the literature in the first three-plus decades post-Pigou.
the inspiration for Coase’s analysis in “The Problem of Social Cost,” took up the subject of road
congestion as little more than an expository vehicle to illustrate a larger point about investment
in industries characterized by increasing costs. Each of these authors recognized that externalities
could cause competitive market outcomes to diverge from the optimum—Ellis and Fellner
(1943, 510) going so far as to call them the only “genuine” type of external diseconomy—but
with the exception of Knight (about whom more below) did little or nothing in the way of
furthering the analysis.

A second indication of the demise of interest in externalities can be found in the major
statements on welfare economics that appeared over this same period, works that one might
expect to have built upon Pigou’s foundations: Abba Lerner’s The Economics of Control:
Principles of Welfare Economics (1944), Melvin Reder’s Studies in the Theory of Welfare
Economics (1947), Paul Samuelson’s Foundations of Economic Analysis (1947), I.M.D. Little’s
Critique of Welfare Economics (1950), Graaff’s (1957) Theoretical Welfare Economics, and
Francis Bator’s “The Simple Analytics of Welfare Maximization” (1957). Lerner, in fact, paid no
attention to external effects of any type in his Economics of Control, a book dedicated to
showing the beneficial effects of a purely competitive system, the workings of which he argued
could be replicated within a planned system. Samuelson (1947, 208), for his part, made passing
reference to Pigou’s discussion of “technological external economies and diseconomies,”
including smoke nuisance, in the lengthy welfare economics chapter of Foundations, but he did
not pursue this line of analysis. He likewise acknowledged the possibility of Veblenesque
interdependence effects on tastes and preferences on the consumption side, but immediately
assumed them away (224). Little (1950, 130), meanwhile, mentioned external effects including
“smoke, noise, and smells” only in passing in his *Critique of Welfare Economics*, and was content to note simply that they work as an impediment to optimality.

Reder went a bit further in his *Studies in the Theory of Welfare Economics* (1947, 62-67), devoting some six pages to an elaboration of external effects in production and consumption as part of his discussion of “obstacles to the attainment of maximum welfare.” Though most of his illustrations went to external effects of the increasing-decreasing returns and Veblen effects varieties, he did instance a factory the smoke from which damages a nearby laundry as one of his illustrations. Nonetheless, he chose to assume away these effects for the meat of his analysis. Bator, writing a decade later, adopted a similar approach, noting the various types of external economies and diseconomies that work as impediments to welfare maximization in a competitive system, but only by way of acknowledging complications that he had assumed away in his analysis. Graaff’s treatment was similarly spartan.

One comes away from these works with the impression that the broad group of divergences between private and social net product pointed to by Pigou were either ignored or, if mentioned, quickly assumed away. Though some authors admitted that these effects posed a problem for the efficient operation of a competitive market system, the precise nature of the problem and what might be done about it was not seriously probed. In short, there was no strong sense given that externalities were the sort of pressing social problem, central to the analysis of economic welfare, suggested by Pigou.

The idea that one could safely ignore external effects in the analysis of competitive equilibrium, though, was not universally shared, and the practice of doing so came in for some criticism. James Meade, then of the London School of Economics (LSE), found it very odd that a book as “brilliant” as Lerner’s could omit any serious discussion of the various categories of
external economies and diseconomies. On the consumption side, Meade pointed to Lerner’s failure to consider both envy effects of the Veblen variety and—more importantly for our purposes here—positive spillovers, such as when a person’s act of beautifying his house “give[s] pleasure to his neighbors as well as himself” (1945, 53). Lerner’s neglect of these issues and their influence on allocation processes within the planning system he advocated, said Meade, had left him wanting on “a basic, if not the basic, problem of the welfare economist” (55, emphasis added). Meade had similar concerns regarding Lerner’s failure to consider external effects associated with production:

There may well be perfect competition all round and yet a misuse of resources because, for example, firms are not charged for the damage which their smoke causes in the district; because they are not charged (or rewarded) for the pain (or pleasure) which the design of their building causes as a part of the surrounding landscape; or because of the many other ways in which they are not charged or paid for the various disadvantages or advantages which their actions may confer on others. (57)

“Here,” said Meade, “is a whole range of effects demanding State control (whether by a system of taxes and subsidies or by other means),” but Lerner had elected to ignore them (57).

Meade was not alone, however. William Baumol, who was also at LSE at the time, leveled a similar charge against Samuelson in his 1949 review of *Foundations* for *Economica*, suggesting that if economists followed Samuelson in assuming that each individual’s utility depends solely on his own consumption and thereby ignored the potential for a wide range of associated external diseconomies, they would “permit some of the most interesting problems of

---

13 M. F. W. Joseph (1944) offered an identical criticism in his *American Economic Review* essay on Lerner’s book. One of the more curious features of the externalities literature between 1940 and 1950 is that a number of the mentions of them in the journals occur in book reviews, with the reviewer taking to task the book’s author for neglecting these important issues in his analysis. See also Bauer (1946, 149) and Vickery (1950, 425).
welfare economics to slip through [their] fingers” (1949, 166). Three years later, Baumol chastised his fellow economists for relegating external effects generally to the category of “freakish exceptions” and for doing so “with little attempt at justification.” In fact, he went so far as to suggest that there was an ideological method to this madness, in that by categorizing external effects as “exceptions,” the “defenders of the existing order” were able to render further discussion them, and thus discussion of the possibilities of government corrective action, “pointless” (1952, 23).14

These criticisms laid down by Meade and Baumol, though, did little to reverse the trend of considering welfare issues sans attention to the variety of external effects dealt with by Pigou. As we shall see, in fact, the eventual renewal of attention to external effects had little to do with filling the perceived lacuna in welfare theory at which these critiques were pointed.15

IV. Externalities and the Competitive Market System

After Pigou, the framing of the discussion of externalities also changed significantly, and the change was almost immediate. On the rare occasions when externalities did enter the picture, it was not as an object of study or as an indication of a social problem to be addressed, as they had been for Pigou and as they increasingly featured in the post-1960 literature, but as a theoretical construct with potential bearing on the efficiency of competitive equilibrium. One cannot understand the “externality theory” of the period without bearing in mind this essential context—

14 Baumol (1952, 23) also offered a critique of Lerner’s analysis that was similar to Meade’s.
15 It may not be coincidental that three of the individuals most strongly emphasizing the importance of external effects in the latter half of the 1940s and the early 1950s—Meade, Baumol, and Myint—were all associated with LSE. Whether this is more than mere coincidence is difficult to say, though Robbins’ lectures on “The Theory of Economic Policy” in 1946-47 and 1947-48 had a Pigovian flavor and took up the subject of external effects. This may have influenced Myint and Baumol, who were students of Robbins at the time. Meade had spent a postgraduate year at Cambridge in the early 1930s, but his interests there were primarily in macroeconomic analysis.
that the focus was not on externalities per se or what should be done about them. In the economics literature, at least, they were not policy-relevant.

A. Externalities, Competition, and Ideal Output

Though Pigou had devoted significant attention to the various types and manifestations of divergence between private and social net products, he had given little attention to the reasons for divergences of the externality type other than to say that, under competitive conditions, it was “difficult to exact payment” for benefits or harms in these situations (1932, 184). It is here that we find the seeds of the “missing market” conception of externalities later elaborated more explicitly and formally by Arrow (1969),\(^\text{16}\) and it was Chicago’s Frank Knight who, in 1924, began to connect the causal dots, turning the discussion away from externalities per se and toward their implications for the efficiency of competitive markets. Knight’s concern was to prop up the competitive system against Pigou’s criticism of it and, specifically, to show that the external economies and diseconomies pointed to by Pigou neither represented failures of competition nor required the application of Pigou’s tax and subsidy remedies.\(^\text{17}\)

Contra Pigou, Knight argued that the source of these problems lay not in the functioning of the competitive market system but in the fact that, in situations of external economies, “The most essential feature of competitive conditions is reversed, the feature namely, of the private ownership of the factors practically significant for production” (1924, 586). Knight illustrated this by drawing on Pigou’s example of road congestion, where a narrow but good road would be overused and a broad but bad road running between the same locations underused.\(^\text{18}\) Pigou had

---
\(^\text{16}\) Papandreou (1994) and Berta (2015), not without some justification, attribute this to Arrow (1969), but the seeds of this conceptualization go back much further.

\(^\text{17}\) That Knight’s concern here was the efficiency of a competitive market system, rather than externalities per se, is reinforced by the fact that half of his article takes on Frank D. Graham’s (1923a; 1923b) recent criticisms, grounded in the analysis of social costs, of free trade doctrine.

\(^\text{18}\) Pigou had used this illustration in both Wealth and Welfare and the first edition of The Economics of Welfare (1920) but eliminated it in subsequent editions, likely in response to Knight’s critique.
suggested that a tax on the use of the good road would be necessary to bring about efficient utilization rates. Knight did not dispute that Pigou’s tax *could be* employed to generate optimal utilization. Instead, he asserted that if the good road was privately owned, users would be forced to pay for the additional benefits conferred by the good road, as a result of which the ideal output “will be brought about through the operation of ordinary economic motives,” just as it would be through Pigou’s proposed tax (1924, 586-87). It bears emphasizing, however, that Knight’s concern here was not so much with congestion per se, but to show that the competitive system can yield optimal outcomes in a wider range of situations than Pigou’s analysis had allowed.\(^{19}\)

This embedding of externalities discussion in the theory of competitive equilibrium transformed Pigou’s difficulties in exacting payment into, as Cambridge economist Richard Kahn put it, an “imperfection of the pricing system” (Kahn 1935, 16), and this view became a staple of the admittedly thin externalities literature. Thus we find Viner, for example, commenting in a passing mention of externalities, set within a discussion of the Graham-Knight debate on trade doctrine, that the “Conceivably important instances of external technological diseconomies” that can be observed in “the grazing, hunting and fishing industries” occur because “no rent is charged for the use of valuable natural opportunities.” This, he said, leads to overexploitation (1932b, 397n.1).\(^{20}\) Kahn likewise attributed smoke externalities to the fact that “the individual does not have to pay for the damage which he does to others,” or, as he also described it, “does not have to pay a price for the air which he utilizes” (1935, 16). So too Ellis and Fellner, who argued that the failure of the market to achieve the optimal outcome was due to “technical or institutional circumstances” in which “scarce goods are treated as though they were

\(^{19}\) Specifically, Knight’s argument was that increasing and decreasing returns under conditions of private ownership may not cause a deviation from the optimum.

\(^{20}\) Viner’s larger concern here was external economies and diseconomies of the increasing and decreasing returns variety.
free,” a problem that they attributed to a “divorce of scarcity from effective ownership” (1943, 511).

Knight, Kahn, and Ellis and Fellner, then, placed resource ownership issues at the root of the pricing system problems that allowed for the externality in the first place. For Kahn and for Ellis and Fellner, at least, this pointed up the differences between the economists’ models of a competitive market system and competition in the real world. In an “ideal world,” Kahn noted, prices would everywhere be equivalent to the marginal cost to society, and thus “every factory, having to pay for the damage which its own smoke caused to others, would emit the ideal quantity of smoke” (1935, 16). But as Ellis and Fellner emphasized, the departures of “actual competition” from this ideal world are “striking” (1943, 511). Externalities thus raised questions about making judgments for the real world based upon the world in the model.21

B. Untangling Externalities

It was only in the 1950s that externalities once again became the object of theorizing, rather than simply the subject of the occasional footnote or passing comment, but the approach here was very different than that found in Pigou and his predecessors. Though the literature remained very thin, there was an effort to fill the lacuna in recent welfare theorizing—lamented by Meade and Baumol—through the analysis of how and why different types of external effect might introduce inefficiencies (toward which Ellis and Fellner had made a halting step in 1943). Along the way, efforts were made to untangle the various classes and situations of divergence between private and social net products set out by Pigou. Papandreou (1994, ch. 2) has characterized the 1950s as a period during which the notion of external economies and diseconomies expanded, but the reality is that Pigou’s broad notion had been carried through in the slim literature of the twenties,

---

21 On economics as a modeling science, see Morgan (2012).
thirties, and forties, even if virtually all of the detailed analysis went to situations of increasing and decreasing returns. What changed during the 1950s was that economists began to seriously examine other types of external effects in the process of differentiating between the various categories and their implications for the efficiency of competitive equilibrium.\footnote{It should be noted that one key facet of this differentiation occurred already in the early 1930s when Viner (1932a) drew out the distinction between “technological” and “pecuniary” economies and diseconomies as part of his analysis of conditions of increasing and decreasing returns.}

\textbf{i. Unpaid Factors and Atmosphere}

Ellis and Fellner, as we have already noted, had distinguished between external diseconomies of the decreasing returns type and those of the negative externality type—the latter inducing inefficiency because of the “divorce of scarcity from effective ownership.” It was almost a full decade, though, before any further attempt was made to disentangle the variety of inefficiency-generating external effects and, in particular, take up externality-type questions of the modern variety. James Meade’s “External Economies and Diseconomies in a Competitive Situation” (1952), was a product of his much larger attempt to bring the tools of welfare theory to bear on the analysis of trade and development, undertaken under the auspices of the Royal Institute of International Affairs.\footnote{See Meade (1955), as well as Meade’s Nobel autobiography (http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1977/meade-bio.html), accessed September 5, 2017.} While this might seem an unusual basis upon which to undertake an analysis of external effects, Meade’s attempt to distinguish between external economies resulting from an “unpaid factor” and those owing to “atmosphere effects” was directly related to questions of international trade and economic development, where external effects, particularly of the increasing and decreasing returns varieties, had figured prominently in the literature.\footnote{On the trade front, the above-cited works by Knight (1924) and Viner (1932) are particularly noteworthy.}

Meade’s article included his now well-known illustration of external effects between beekeepers and apple orchard owners, the culprit being what Meade labeled “unpaid factors of
production” (1952, 56ff): The apple farmer cannot charge the beekeeper for the input he provides to the production of honey. Meade was the first to formalize this class of external effects, and he did so using a two-industries competitive model that became the basis for much of the subsequent theorizing about externalities. In the most basic case described by Meade, the farmer’s apple blossoms provide food for the bees, and Meade modeled this relationship as follows:

\[ x_1 = H_1(l, c, x_2), \]

where \( x_1 \) and \( x_2 \) represent outputs in beekeeping and apple farming, respectively, and \( l \) and \( c \) are labor and capital inputs. Here, \( x_2 \) represents the “unpaid factor,” which thus earns a return less than its marginal social product. Meanwhile, inputs in the beekeeping industry receive a return greater than their marginal products (57). The result, of course, is an inefficient allocation of resources to each of these production activities.

Meade then turned his attention to external effects related to the creation of “atmosphere,” which he ascribed to “the fact that the activities of one group of producers may provide an atmosphere which is favourable or unfavourable to the activities of another group of producers” and is unaffected by the scale of the affected industry (1952, 62). Meade instanced a situation in which afforestation efforts in a particular area increase rainfall and so benefit wheat production in that area, and he formalized the relationship as follows:

\[ x_1 = H_1(l, c_1)A_1(x_2), \]

---

25 Meade also explored two-way external effects here, allowing that the bees may provide fertilization services for the farmer’s trees. We take up only the one-way case for the sake of brevity.

26 In the unpaid factor case, in contrast, the benefit to the affected firms of a given output by the industry controlling the unpaid factor diminishes as the size of the affected industry increases.
where $x_1$ is wheat output and $x_2$ is output in the timber industry, with the effect of timber output on wheat production being a function of the relevant atmosphere created, $A$. Inputs into wheat are paid at the value of their marginal social product, while those used in timber production are paid a value less than their marginal social product (63).

What was to be gained by differentiating between these two types of external effect—the unpaid factor on the one hand and the atmosphere effect on the other? Meade’s analysis demonstrated that these two types of effect influence output levels in different ways, with implications for potential efficiency-generating remedies. The unpaid factor case involves constant returns for society as a whole but not for each industry, while in the case of atmosphere effects there are constant returns for each industry but not for society as a whole (56, 67). The implication that Meade drew from this was that the efficient input mix will obtain in the unpaid factor case only if there are in place appropriate taxes (on bee keepers) and subsidies (to apple farmers) that bring private and net social products into line, and that the revenue from the taxes would be precisely the amount required to pay the corresponding subsidies (57). In the afforestation case, in contrast, if each factor is paid its marginal social net product, those payments will exceed industry revenue. As such, the subsidies necessary to bring private and social net products into line must be financed out of general tax revenues (1952, 62). The sub-optimal allocations generated by these different forms of external economy, then, have very

---

27 In the unpaid factor case, doubling inputs to apple farming doubles apple output, but doubling inputs to beekeeping does not double honey output unless apple farmers also double their output to provide the necessary additional food for the bees. In the atmosphere case, in contrast, doubling wheat inputs will lead to a doubling of wheat output with timber production held constant, but doubling all inputs to both timber and wheat will more than double wheat production due to the positive atmosphere creation associated with the additional timber output.

28 The question of whether a combination of taxes and subsidies is required in the unpaid factor case became a point of some contention in the subsequent literature. See, e.g., Baumol (1972).

29 Of course, if the external effect were a negative one, the taxes needed to equate private and social net products would work as an addition to the public purse.
different economic underpinnings and necessitate different remedies in order to ensure efficient competitive equilibrium output levels.

ii. Varieties of “Direct Interdependence”

Meade (1952, 67) acknowledged that his analysis provided at best a partial clarification of the muddle that was the theory of external economies and diseconomies. Indeed, such was the state of the discussion that Stanford’s Tibor Scitovsky could note two years later that the existing definitions of external economies were “few and unsatisfactory” and that it was “nowhere made clear” exactly what types of activities properly fall under this heading. But this was only part of the problem, Scitovsky said. There was not even agreement on their relevance. Some commentaries suggested that these phenomena were “exceptional and unimportant,” while others implied that they were “important and ubiquitous” (1954, 143). A further measure of clarification was thus in order.

Scitovsky located the heart of the external effects problem in the existence of a “direct interdependence” between agents, where by “direct” he meant interdependencies that do not “operate through the market mechanism” (144). Absent these interdependencies, he said, equilibrium in a competitive economy will be Pareto optimal. This emphasis on an inadequacy in the pricing mechanism, of course, was of a piece with most earlier commentaries, and Scitovsky’s attachment of agent interdependence to this was drawn directly from Meade’s mathematical statement of the problem.30 While interdependencies were endemic to economic relationships, what made these external effects inefficient was the fact that they were not transmitted through the pricing system.

---

30 Meade himself had not discussed interdependence as a defining feature of the externality issue, nor did he use the term.
Scitovsky then proceeded to distinguish between four categories of “direct interdependence” that may exist within a competitive system, aligning them with illustrations that had been carried through the earlier literature. The first of these involve situations in which one consumer’s satisfaction is a function of the satisfaction of other consumers, often referred to as “Veblen effects.” These, Scitovsky said, are “undoubtedly important,” and fed into the ongoing controversy in welfare economics and economists’ hesitancy to make consumer-related welfare statements. The second class of direct interdependencies identified by Scitovsky was producer actions, such as the generation of smoke or noise, that influence consumer satisfaction in ways not channeled through the pricing mechanism. But these Scitovsky considered “exceptional” because, as he put it, they “can be and usually are eliminated” by regulations of various types. Third, he said, a firm’s output may be influenced by the actions of other agents in ways not related to their production and consumption activities—e.g., as the result of a new invention that is freely available. These effects, Scitovsky contended, are rendered “unimportant” by the existence of a patent system. Finally, the output of one firm may be impacted by the production activities of other firms, as in Meade’s illustration of the apple farmer and the bee keeper. But here, too, Scitovsky found the external effect “unimportant,” largely because “examples of it seem to be few and exceptional”—evidence for which, he said, could be seen in Meade’s use of the “somewhat bucolic” examples of bees, orchards, and timber. Meade’s choice of illustrations, said Scitovsky, was “no accident,” as examples from industry are extremely rare (1954, 144-45).

---

31 Pigou (1928, 185) had previously made this point.
32 Scitovsky found that he could identify only two examples that fit his definition—that in which a firm benefits from a labor market developed through the existence of other firms, and resources that can be used at no cost to the individual agent but are limited in supply. Scitovsky closed this discussion by referring the skeptical reader to Meade’s 1952 article, which he believed would “convince [the reader] of the scarcity of technological external economies” (1954, 145). It bears mentioning that Meade’s example of the bee keeper and the orchard owner has real-world currency though, apparently unbeknownst to Meade, there are numerous examples of private contracting
Thus, both the existing literature on the subject and a bit of reflection suggested to Scitovsky that, save for Veblen-type effects on tastes and preferences, the theory of competitive markets did not need to concern itself with external economies and diseconomies, including those interdependencies which now go by the name “externalities”—the second and fourth classes that he had delineated. The fact that these interdependencies were either “exceptional” or “unimportant” meant that they could be safely assumed away.

iii. Probing Market Failure’s Anatomy

It was Francis Bator’s elaboration of “The Simple Analytics of Welfare Maximization” in 1957 that introduced the term “externality” into the economics literature (1957, 42, 43). There is some irony in this, in that Bator echoed Scitovsky’s sentiment that these effects were “unimportant” (1957, 42), based upon which he discussed the welfare implications of competitive equilibrium sans attention to external effects. It was his 1958 follow-up article, “The Anatomy of Market Failure,” though, that gave form to the concept of externality as we know it today.

between bee keepers and orchard owners that account for the externalities involved. See, e.g., the discussions in Cheung (1973) and Johnson (1973).

33 Scitovsky’s attitude here provides additional evidence for Baumol’s (1952, 23) contention, noted above, that there was a tendency “to relegate cases where competitive action is not conducive to social welfare to the category of freakish exceptions.”

34 Bator first used the term in his 1956 MIT Ph.D. thesis, from which his 1957 and 1958 articles were derived (Bator 1956). Whether it was Bator who coined the term is hard to say. Paul Samuelson, Bator’s MIT colleague, used it at nearly the same time in an article on the subject of intertemporal price equilibrium that appeared in Weltwirtschaftliches Archiv (Samuelson 1957). Samuelson noted simply that “knowledge is a resource loaded with externality” (1957, 210). Bator’s article seems to have appeared first and was certainly more widely read, being cited nearly fifty times over the next dozen years as against eighteen cites to Samuelson’s piece. (Source: Google Scholar, accessed August 22, 2017). Moreover, none of the citations to Samuelson’s article were in the welfare economics/externalities context. There is no input from Samuelson acknowledged in Bator’s obligatory opening footnote, nor is mention of Bator made in Samuelson’s article. The term was picked up relatively quickly though, being used another half-dozen times before 1960. It was also Bator who introduced the term “market failure” into the literature, this in his 1958 article, “The Anatomy of Market Failure.”
The initial location of the modern concept of externality in the theory of competitive equilibrium rather than in real-world phenomena is perhaps nowhere better exemplified in the literature than in Bator’s opening paragraph:

What is it we mean by “market failure”? Typically, at least in allocation theory, we mean the failure of a more or less idealized system of price-market institutions to sustain “desirable” activities or to estop “undesirable” activities. The desirability of an activity, in turn, is evaluated relative to the solution values of some explicit or implied maximum-welfare problem. (1958, 351)

It was the role played by externalities in this theoretical world, and not “the efficiency of ‘real life’ market institutions,” that Bator set out to explore (352).

Bator, like Scitovsky, found the extant externalities literature “rich but confusing” (1958, 356), and he was not convinced that Scitovsky’s efforts had done much in the way of furthering economists’ understanding of the concept. Scitovsky’s notion of “direct interaction,” Bator said, “begs more question than it answers,” as it does not get to the question of why this unpriced interaction occurs in the first place. He also considered Ellis and Fellner’s emphasis on the divorce of scarcity from effective ownership “misplaced” because it could not explain why certain types of goods, such as the services of a bridge, are not priced in the marketplace given that exclusion is clearly possible (352, 361). A new, or at least enhanced, conceptual framework was thus in order.

To get around the limitations he had identified in the existing literature and create a workable concept of market failure, or “statical externalities,” Bator identified three “polar types” of failure. The first, “ownership externalities,” was essentially equivalent to Meade’s “unpaid factor” case, and here Bator found Ellis and Fellner’s “divorce of scarcity from effective
ownership” to be “the binding consideration,” preventing certain resources from being priced in the market (364). Examples given by Bator include Meade’s bees and apple blossoms, fisheries, mineral extraction, tenancy investment, smoke pollution, and labor force skills training—that is, situations that tend to be classified within the boundaries of externalities as we think of them today.

The second category delineated by Bator was “technical externalities,” where indivisibilities or increasing returns to scale give rise to non-convexities in the production set (365-69). These, he argued, are “much more important” than externalities of the first type (365), and unlike with that class of externalities, appropriability does not resolve the problem. For example, if the marginal cost of bridge crossings is zero, marginal-cost pricing by a private owner is not sustainable, while any positive price results in an inefficiently low number of crossings. Nonconvexities, then, explained a class of external effects that would preclude the attainment of the optimum even in the absence of nonappropriability issues.

Finally, said Bator, market failure may arise due to “public good externalities” of the type described by Samuelson (1954; 1955). While Samuelson had made a passing reference to the external economies that attend joint consumption (1954, 389), his concern was with “the theory of optimal public expenditure” (387). Bator’s focus, in contrast, was on why the market will fail to supply these goods efficiently in the first place, and he identified their public, joint-consumption nature as the reason why it was impossible for the market to generate the efficient price-output mix (1958, 369-71). The value of the good enters jointly (and positively or

35 Bator also included Meade’s atmosphere effects in this category, whereas Scitovsky had considered them of the same form as the unpaid factor—Bator’s first category.
negatively) into the utilities of multiple individuals, according to Bator, because “my party is my neighbor's disturbance, your nice garden is any passerby's nice view, my children's education is your children's good company, my Strategic Air Command is your Strategic Air Command, etc.” (370). Even absent problems of nonappropriability and nonconvexity, the prices necessary to induce optimal production by the agents whose actions are attended by this class of externality will not themselves generate optimal allocations because of the incentives facing the agents involved.

Bator’s classification scheme brought some additional clarity to the externalities concept, teasing out fundamental differences in these types of external effect and showing how and why each caused competitive equilibrium outcomes to diverge from the optimum. That said, his analysis by no means resolved the muddle, as we can see from the many successive attempts to refine the concept. 37 Bator himself acknowledged that some phenomena are, in reality, “blends” of the three types that he had identified, 38 but he was of the mind that this three-fold demarcation was analytically helpful (376). In time, of course, it was the first of these, “ownership externalities,” that came to define the term, “externality,” in the economics literature. The others, meanwhile, eventually were classified as separate instances of “market failure,” alongside externalities, monopoly, and so forth.

* * *

What we observe in the literature in the four decades following the publication of Pigou’s treatise, then, is not an expansion of the notion of externalities, as Papandreou (1994) has suggested, but instead a delineation of categories within the broad boundaries, and including the wide range of activities, originally laid out by Pigou—and, indeed, by Mill and Sidgwick in the

37 See Papandreou (1994) for a discussion.
38 He cited the lighthouse as an example.
nineteenth century. What changed along the way was not the scope of the externality concept but its very nature. To the extent that they were discussed in the literature, externalities were not characterized as real, policy-relevant phenomena. Instead, they were depicted as theoretical relationships that interfered with the ability of a competitive market system to satisfy the dictates of optimality. What remains is provide an explanation for this.

V. Explaining the Lacuna and the Muddle

One possible explanation for the deemphasis on externalities, post-Pigou, is that external effects of the type that concern us here were considered empirically unimportant, a hypothesis for which we certainly find evidence in the theoretical literature. Viner, for example, suggested that inter-firm external effects “can be theoretically conceived, but it is hard to find convincing illustrations” (1932a, 41), an assessment echoed by both Scitovsky and Bator. As multiple authors pointed out, the examples given did not go to real-world problems of external effects. The factory whose smoke impacts a neighboring laundry or area is a far cry from large-scale pollution externalities, and many other examples were even less connected to what might be considered significant real-world phenomena. Moreover, such external effects as might otherwise exist, it was said, likely had already been dealt with via the legal-political system. They were not, then, a problem with which economists needed to concern themselves.

But this is at best a partial explanation, if for no other reason than that it does not account for the acknowledged muddled state of the discussion. To fully comprehend the situation, we must examine the transformations taking place within and beyond economics during and immediately after the interwar period. First, of course, there is the Great Depression and the

---

39 Viner mentioned road congestion as “one possible instance” of such effects (1932a, 41).
40 Bear in mind that in the analysis of Pigou and his predecessors, we find no evidence for a perceived muddle.
wartime planning that followed. Together, these crises preoccupied the attention of economists for some fifteen years and provided a contextual backdrop against which externality phenomena would seem to pale in comparison. Second, the tradition in which Mill, Sidgwick, Marshall, and Pigou had worked was one in which economics was focused on dynamic issues of development and growth, and the external effects pointed to by those writing in the decades following Pigou were very much the byproducts of the growth process that framed these earlier works. As Lionel Robbins argued in his *Essay on the Nature and Significance of Economic Science* (1932), however, this was not the approach that increasingly was coming to dominate economic analysis. Instead, the emphasis was on choices made under the influence of scarcity. This move brought with it a static emphasis, including a focus on the properties of equilibrium in alternative market contexts, and was reinforced by the technical tools that were being brought to bear on the analysis of economic problems. The influence of the socialist calculation debate looms large here (e.g., Lerner 1944), as it at once called into question the ability of a competitive market system to generate an efficient allocation of resources and stimulated a significant push—by friends and foes alike—to elaborate the theoretical conditions necessary for the attainment of that optimum. One byproduct of these efforts was that “externalities” were pushed to the side during this period. A second is that, when they were taken up, it was in the context of this new-found preoccupation with the properties of competitive equilibrium. Evidence for this assessment can be found both in the nature of the discussion of externalities in section IV and from certain comments made by those contributing to that literature, and it features prominently in explanations for both the lacuna and the attempts to clarify the externalities muddle.

41 That is, the implications of increasing and decreasing returns were still very much part of the conversation, even if externalities were not.

42 On this point, see Backhouse and Medema (2009).
Several of the individuals involved in these discussions expressed a belief that incorporating external effects into the analysis would interfere with the elegance of the theorizing process—a view that, admittedly, may have been influenced by perceptions of empirical unimportance. Both Reder and Little, for example, were explicit in expressing their comfort at leaving external effects to the side in their analysis. Reder did not believe that “this assumption greatly restricts the validity of our analysis” (1947, 67), and Little complained that the inclusion of these effects “destroy[ed]” its “precision” (1950, 130). Neither, it should be noted, said anything about remedies for these external effects, being content simply to note that they could cause market outcomes to deviate from the optimum.

Perhaps the most elegant and insightful statement of what was at stake here, though, was provided by Myint, who made this abstract, formal approach to welfare theorizing turn on economists’ newfound preoccupation with the Pareto optimality:

The Paretian theory of the Optimum is concerned with the basic methodological problems of welfare economics rather than with its practical application. Its main purpose is to show that it is possible to formulate a stringently demonstrable concept of the Optimum which avoids the traditional bugbears of welfare economics, viz. interpersonal comparisons of utility and value judgments. In order to achieve its aim, however, it has to sacrifice realism and assume the ideal conditions of Perfect Competition with perfect mobility and divisibility of factors and perfect knowledge. The fact that Perfect Competition could never be attained in practice even under the most favourable conditions does not invalidate the concept of the Optimum. (1948, 187-88)\(^{43}\)

\(^{43}\) Pigou’s approach, said Myint, was more realistic, but it ran into the interpersonal welfare comparisons “which make his analysis unacceptable to the purist economist” (1948, 188).
This turn in welfare economics was, for Myint, a mixed blessing. On the one hand, he said, “It has been a great achievement to formulate a stringent concept of the Optimum; it is the necessary foundation of a scientific welfare economics and without it we cannot conceive the Deviations from the Optimum.” But yet, he continued, because “the normal feature of economic life consists in the deviations from, and not the attainment of, the Optimum,” the current line of analysis “is only a beginning of welfare economics.” Analysis of actual deviations, Myint contended, would required “a more realistic model than that of Perfect Competition which virtually assumes away all possible frictions and faults” (188).

Bator’s attitude toward external effects, expressed a decade later in his “Simple Analytics of Welfare Maximization” (1957), only adds weight to our explanation. Bator admitted that he had derived his results using “the simplest statical and stationary neoclassical model,” an approach that he justified on the grounds that introducing additional complexity for the most part would not vitiate his results. However, he did allow that there was “one kind of complication” which would—allowing for “(nonpecuniary) external economies or diseconomies of production and consumption” (1957, 42).44 Bator’s response, though, was not to include these complications, but to defend his decision to exclude them from his analysis.

Questions of empirical relevance appear to have factored into Bator’s attitude. He, like Scitovsky several years earlier, argued that the “very pastoral quality” of the beekeeper and orchard owner example “suggests that in a statical context such direct interaction among producers—interaction that is not reflected by prices—is probably rare” (Bator 1957, 42, emphasis added). Bator was also rather unconcerned that his model had ruled out consumer-side

---

44 Here, Bator (1957, 42) indicated that Meade’s illustration of the beekeeper and apple farmer was the “stock example,” which is interesting given that Meade’s article had been published only five years earlier and had been cited fewer than 10 times in the literature to that point. This may be an indication of an “oral tradition” in externalities to which Coase (1960, 39) referred three years later.
external effects—including “such phenomena as Y tossing in sleepless fury due to X’s ‘consumption’ of midnight television shows; or X’s temperance sensibilities being outraged by Y’s quiet and solitary consumption of Scotch”—and the impact of producer decisions on consumer welfare—such as when “Y’s wife [is] driven frantic by factory soot,” or “X [is] irritated by an ‘efficiently’ located factory spoiling his view” (43).

Bator acknowledged that it may be possible to include these various types of interaction effects into a formal model such as the one he was considering and even that they had real-world relevance. The “neighborhood” phenomena to which he had referred, Bator noted, are far from “illusory.” Yet, he was of the mind that “it is not very fruitful to take account of them in a formal maximizing setup,” as doing so likely would come only at significant cost and the welfare results derived would lose much of their specificity and meaning (43). Because of these complications, he said, “most formal models rule out such phenomena,” a practice that he defended:

There is no doubt that by so doing they abstract from some important aspects of reality. But theorizing consists in just such abstraction; no theory attempts to exhaust all of reality. The question of what kinds of very real complications to introduce into a formal maximizing setup has answers only in terms of the strategy of theorizing or in terms of the requirements of particular and concrete problems. For many purposes it is useful and interesting to explore the implications of maximizing in a “world” where no such direct interactions exist. (44)

45 Bator considered public goods an exception to this, citing Samuelson’s “original and definitive treatment” (1957, 43n.43, 44n.44). He apparently did not believe that other types of external effects could be modeled in a way that generated such definitive results.

46 See also Fisher (1956, 416n.6), who offers a similar perspective. Stephen Sosnick, in contrast, pronounced himself to be unconcerned with the theoretical properties of perfect competition, but instead with the “theory of workable competition,” which he defined as “an attempt to indicate what practically attainable states of affairs are socially desirable in individual capitalistic markets” (1958, 380). Sosnick cited external economies and diseconomies as two of a number of factors that illustrate “that the perfectly competitive structure and conduct are unattainable in any real market” (384).
This statement is particularly instructive as to the attitude of the day, coming as it did from a participant in the game rather than from, say, a more critical observer such as Myint. Economic theorists were busying themselves exploring the properties of competitive equilibrium and felt it necessary to abstract from various features of reality to build tractable models of the perfectly competitive market process. The analysis of external effects was sacrificed to the needs of the theorizing process.

How, then, do we explain the move by economists to begin to theorize about externalities? A similar set of larger professional forces play a role here. External economies and diseconomies—including externalities, to the extent that they were mentioned—were seen as an impediment to the efficiency of competitive equilibrium. It eventually became necessary to assess how any why this was the case, and it thus seems only natural that economists would attempt to model these effects in a static equilibrium context in order to ascertain more precisely the reasons underlying the resulting inefficiencies. Doing so, however, meant teasing out the distinctions between various categories of external effects in order to model them properly and so get at the economic logic underlying the inefficiency that each engendered.

But there is more to the story here. We must also bear in mind that external effects of the increasing and decreasing returns variety had a significant presence in the international trade and development literature throughout this period. Though there was a significant dynamic element to these problems, economists increasingly treated them using static welfare analysis and this, too, stimulated efforts to tease out the distinctions between various types of external effects. Indeed, Meade, as we have already noted, took up his analysis of unpaid factors and atmosphere effects as part of an attempt to apply welfare theory to trade and development issues—though Scitovsky lumped Meade’s analysis into the competitive equilibrium category—and Bator, too,
was motivated in party by issues related to the economic growth process, as evidenced in his Ph.D. Thesis, from which the 1957 and 1958 articles were drawn. Scitovsky (1954), for one, was critical of the application of static analysis to problems that he saw as inherently dynamic and argued that this had contributed to the muddled situation in the analysis of external effects. Bator, though, saw no need to pull back from static competitive analysis in such work, and his defense of this approach reinforces our assessment that the competitive modeling turn is central to the explanation for how externality analysis evolved in the decades following the publication of Pigou’s treatise:

whether in a statical or dynamical context, the questions are all relevant to whether a decentralized price market “game”—perhaps “for real” by genuine profit seekers, or by socialist civil servant plant managers following an injunction to “maximize profits,” or, perhaps for no less “real,” by technicians following out a computing routine—will or will not sustain a Pareto-efficient configuration once the shadow-prices associated with that configuration are specified. (1956, 69)

The nature of equilibrium, rather than the nature of externalities and what to do about them, was the crux of the problem that needed solving.

* * *

The focus on static, competitive equilibrium analysis, then, explains both the lacuna and the reemergence of externalities within welfare theorizing—as well as the form, so different from Pigou’s—taken by this new line of analysis. Even so, it would be difficult to conclude that, by the end of the 1950s, economists had a well-developed theory of external effects, or even that a

---

47 See, esp., Bator (1956, 70-71). Scitovsky attributed the two-headed approach in part to the “separation of the different branches of economic theory,” one of which was “the theory of industrialization in underdeveloped countries” (1954, 143).
great deal of effort had been devoted to trying to work out such a theory. Whatever the reasons for this—perceived empirical irrelevance, letting the tools and modeling strategies set the agenda for research, the desire to come to grips with the competitive Pareto optimum—it is difficult to avoid the conclusion that the discussion was little more than a confused muddle of different ideas about interdependencies among agents and that economists had not yet managed to, or perhaps felt the urge to, attempt to tear apart the various strands and break things down to the essentials—though a few steps were taken along the way by, e.g., Meade, Scitovsky, and Bator. Virtually no attention had been paid to the phenomena themselves; instead, the focus was on the characteristics of these external effects that influenced the attainment of the Pareto optimum. Moreover, there was nothing about remedies for these effects beyond the occasional mention that taxes or subsidies could bring private and social benefits and costs into line. In the mid-1950s, though, we began to see evidence from other quarters that the situation was changing.

VI. Rebirth: Externalities Take Form

The relative thinness of the externality theory literature in the three-plus decades post-Pigou was more than matched by the lack of effort expended exploring specific problems of external effects pointed to by Pigou and potential remedies for them. As we move into the second half of the 1950s, however, we find the seeds of a renewed emphasis on externality phenomena and on how they might best be resolved. Its location, though, was not so much in the realms of high theory as in the emerging subfields of applied economics.

Papandreou (1994, 44-47, 69ff) traces the association of externalities with specific real phenomena, such as pollution, to the rise environmental economics in the 1960s and suggests that this view was solidified in the professional mind by Meade (1973) and Baumol and Oates.
Papandreou is absolutely correct in linking the modern phenomenological approach to the rise of environmental economics, but this conception of externalities was by no means new. Mill, Sidgwick, and Pigou each had taken a phenomenological approach to the subject; as we have seen, however, their concerns were not carried through in the next generation of external effects scholarship. In the 1950s—roughly a decade prior to the time identified by Papandreou—though, things began to change.

Meade (1945) and Baumol (1949; 1952) had expressed concerns about the reality of externalities and the need for economists to confront them in their analysis, but neither made any significant efforts—at that time, at least—to develop this line of research. Indeed, prior to the mid-1950s, such references as were made to externality phenomena that might exist in the real world were of the passing sort and did not move beyond canned invocations of polluting factories, fisheries, and so on—a list to which was added Meade’s beekeeper-apple farmer illustration in 1952. Simply put, there was no economic analysis of externality phenomena. As we move through the latter half of the 1950s, however, we begin to see the concept and analysis of externalities applied to a small set of policy-relevant issues, including road congestion, fisheries, urban renewal, water supply, agricultural tenancy, broadcast frequency allocation and, of course, pollution.

The driving forces behind these various efforts were two. The first, and likely most important, was the perception, not unlike that found in Mill (1848), of a set of pressing social problems requiring a response. The second—in part derivative of the first—was the rise of

\[48\] The textbook literature provides an excellent additional illustration of this transition in action, as the discussion of externalities migrated from chapters on cost theory to self-contained chapters in policy- and applications-oriented sections of the textbooks.

\[49\] Baumol later became actively engaged in the economics of the environment. See, e.g., Baumol and Oates (1971), Baumol (1972), and Baumol and Oates (1975).
“applied” economic analysis (Backhouse and Biddle 2000), including the fields of environmental and urban economics. Perhaps the most significant institutional impetus for environmental economics research came through the founding of Resources for the Future (RFF) in 1952. Public concerns about natural resources shortages—stimulated in part by the significant resources consumption associated with WWII and the Korean War—led President Truman to establish the “Materials Policy Commission” (also known as the Paley Commission) in 1951. The Commission’s report recommended the establishment of a permanent, independent organization the purpose of which was to analyze the country’s natural resource supplies, and RFF was founded, with funding from the Ford Foundation, “to support the conservation, development, and use of natural resources.”50 As “the first think tank devoted exclusively to natural resource and environmental issues,”51 it supported work by economists on natural resource and environmental concerns.

A rich history of environmental economics has yet to be written,52 and delving deeply into this history goes well beyond the scope of the present paper. What is relevant for present purposes is that the literature of the 1950s evidences a flicker of interest in environmental topics and the beginnings of a separation of environmental economics from natural resources economics.53 This uptick in the economic analysis of environmental issues during the latter part of the 1950s was followed by a surge in the 1960s, which saw the publication of highly influential works by Hirshleifer et al. (1960), Kneese (1964) and Krutilla (1967) that helped to

50 Source: http://www.rff.org/about/rff-s-legacy.
51 Ibid.
52 In particular, little work has been done to date on the “institutional” history of the subject and the professionalization of the field. For some general intellectual histories of environmental economics, see Spash (1999), Crocker (2002), Pearce (2002), and Sandmo (2015). The work of Spencer Banzhaf is particularly instructive on particular aspects of this history. See, e.g., Banzhaf (2016); (2017).
53 Representative works from the 1950s include Ciriacy-Wantrup (1952), Gordon (1954), Krutilla and Eckstein (1958), and Eckstein (1958). Each of these was influential in the development of further scholarship in this area.
set the field on its course. Further markers of the growth of this emerging field came on the dissemination front, with the founding of the *Natural Resources Journal* in 1961 and the move by *Land Economics* to focus more heavily on environmental economics beginning in the early 1960s (Spash 1999, 418).

It is almost trivial, then, to make the connection to the rebirth of externality analysis. Environmental problems can be thought of problems of externality, so it is only natural that the increasing concern with environmental problems would take economists back to externality analysis. But this interpretation is in some ways *too* trivial. It is certainly true that pollution had for more than a century been discussed as an illustration of what came to be called “externalities,” and that the extant theory of externalities was *grafted into* the emerging field of environmental economics—perhaps the classic early statement coming from Kneese (1964). But it did not need to be so. Economists taking up environmental issues could have elected to develop a new theoretical framework, adapted to the complex nature of environmental issues, upon which to base their analysis. But they did not, electing instead to pull Pigou’s concept of external effects, as refined to some extent by subsequent commentators, off the shelf to serve as the basis for analysis. This decision, too, is a reflection of extent to which the static optimization approach had come to dominate thinking: Some of the same forces that explain the lacuna and the muddle also explain the form in which phenomenological externality analysis was reborn. Ironically, it was only a bit more than a decade later that we find Kneese (1971) lamenting that the decision to ground environmental economics in the theory of externalities had provided to be a significant impediment to the field’s progress, with the two-agent static models that were standard in the literature ill equipped to deal with the complexity of the phenomena to which they were being applied.
These forces combined to initiate three significant shifts in the analysis of external
effects.\textsuperscript{54} First, while most of the theoretical discussion of external economies and diseconomies
over the previous three decades had emphasized the former—derivative of the increasing returns
spillovers that preoccupied economists taking up these effects—the emphasis in this newly
developing literature was on external diseconomies, of the form now classed as negative
externalities.\textsuperscript{55} Second, the reality of these external effects and of the resulting resource
misallocations was emphasized throughout; that is, externalities were no longer considered, as
Scitovsky had classified them, “exceptional and unimportant.” The third feature that set this
literature apart was the sense that these were problems that merited addressing on the policy
front and that the economist had something useful to contribute to the discussion—as a result of
which we find the authors focusing on externality remedies.

To the extent that remedies had been discussed prior to the mid-1950s, the default had
been to Pigou’s taxes and subsidies as price-related tools which could restore the marginal
equivalences dictated by optimality theory. During the latter half of 1950s, however, these
Pigovian instruments lost whatever small pride of place they had as economists explored the
question of how best to deal with particular situations of externality. What stands out most
vividly here is the lack of any settled sense for how best to address these problems. As one might
expect, Pigou’s discussion of direct state action was the starting point for the discussions; but
there was little confidence expressed in Pigovian remedies. While Little (1957, 155) offered

\textsuperscript{54} This is not to say that the (non-phenomenological) analysis of the implications of externalities for competitive
equilibrium disappeared—far from it, as evidenced by the work of Bator and Graaff (1957). Indeed, the
conceptualization of externalities and the modeling strategies employed in applied externality analysis were adopted
from this earlier literature. But going forward, the analysis of the externality phenomena themselves would proceed
alongside the competitive equilibrium analysis.

\textsuperscript{55} In fact, one could argue that Meade’s decision to explore external economies, rather than diseconomies, in his
analysis of unpaid factors and atmosphere effects was derivative of the focus on external economies in the trade-
development literature.
passing support for Pigovian measures,\(^{56}\) Kahn (1935, 16), Walters (1954, 143), Baumol (1952, 167), and Myint (1948, 192) were far less confident of their efficacy, and for a variety of reasons.\(^{57}\) Other policy options were very much in play. Single owner solutions, akin to that originally developed by Knight, were advocated, with various degrees support, by Gordon (1954), Scott (1955), Buchanan (1956), and Bailey (1959, 288). Even negotiated solutions, often attributed to Coase (1960), featured in the analysis, being given some measure of credence by Bailey (1954, 50-51), Milliman (1956; 1959), Krutilla and Eckstein (1958, 1684), and Turvey (1957, 95-96). The common thread here is that all of these measures were seen to have significant pluses and minuses. On the remedies front too, then, the literature provides us with little evidence of an entrenched Pigovian tradition.

VII. Conclusion

The history of the modern concept of externality is tied closely to both the larger social-economic contexts within which economists lived and worked, as well as the trend of economic thinking. Born of a mid-nineteenth-century concern that the invisible hand was not performing the functions that some had ascribed to it, the externality became, along with other instances of divergence between private and social net product, a core component of Pigou’s 1920 elaboration of instances in which the market, left to its own devices, would not maximize welfare for...

\(^{56}\) Meade (1952) and Graaff (1957) discussed the possibilities of taxes for restoring marginal equivalences in theory, but went no further than that.

\(^{57}\) Myint, for example, writes that though taxes on polluting factories of the sort advocated by Pigou may “ease the situation, a complete remedy of the evil would seem to require the abandonment of the existing pattern of land utilisation and the introduction of a more rational pattern where all such harmful juxtaposition of industrial and residential sites are prevented as much as possible” (1947, 180). Myint also questioned the possibility of accurately monetizing costs such as pollution damage, as well as whether policy makers could even get away with making such calculations if possible. “[W]ho,” he asked, “would dare to assess these evils in money terms?” (184).
and thus where the State could potentially have a role to play in improving economic performance.

Externalities, though, largely disappeared from economic analysis for nearly three decades following the publication of Pigou’s welfare treatise. Such discussions as did occur prior to the late 1950s were bound up in questions of the efficiency of the competitive market system, the implications of the broad class of external economies and diseconomies elaborated by Pigou for optimality, questions of whether the different forms of external economies and diseconomies had differing efficiency implications, and how to model these various forms of economic interdependence. Externalities themselves were considered, as Scitovsky put it, “exceptional and unimportant” and so largely irrelevant to the concerns of the economist.

It was only in the latter half of the 1950s that economists once again began to turn their attention to externalities per se—that is, to the analysis of externalities as real economic phenomena worthy of examination in their own right. The concerns driving this turn in the analysis were largely those that had preoccupied Pigou, and even his predecessors—pollution, overexploitation of natural resources, congestion, and related problems of economic growth and development. The notion of externality came to be the lens through which economists viewed these problems and set to the task of proposing policy remedies. In the process, the scope of economic analysis was broadened to include environmental concerns, making this perhaps the earliest instance of what has come to be called “economics imperialism.”

The attacks launched by Coase, Buchanan, and others against the “Pigovian tradition” in the early 1960s thus were not so much attacks on a straw man as on a man not evident in the economics literature—though perhaps he could be found in economics department hallway conversations. Coase’s discussion of negotiated solutions to externalities in “The Problem of
Social Cost,” then, must be viewed in a different light. It was neither the first suggestion of the possibility of negotiated or private solutions nor the catalyst for the explosion in the externalities literature that followed over the next decade—a decade during which there were more than 450 articles discussing “externality” or “externalities” in JSTOR journals alone. Instead, Coase’s analysis, itself motivated by concerns of a phenomenological sort—the allocation of broadcast frequencies—was caught up in the larger professional interest in externalities that arose in response to perceived problems of industrial pollution, natural resource depletion, and urbanization.

References


58 Source: jstor.org, accessed September 5, 2017. The reaction to Coase’s analysis in “The Problem of Social Cost” is an artifact of its intersection with both the rebirth of the phenomenological concern over externalities and the prior tendency to discuss externalities in terms of implications for the efficiency of competitive equilibrium. Consideration of this issue, though, goes beyond the scope of the present paper.


