

# Algebraic Production Functions and Their Uses Before Cobb-Douglas

---

Thomas M. Humphrey

**F**undamental to economic analysis is the idea of a production function. It and its allied concept, the utility function, form the twin pillars of neoclassical economics. Written

$$P = f(L, C, T \dots),$$

the production function relates total product  $P$  to the labor  $L$ , capital  $C$ , land  $T$  (terrain), and other inputs that combine to produce it. The function expresses a technological relationship. It describes the maximum output obtainable, at the existing state of technological knowledge, from given amounts of factor inputs. Put differently, a production function is simply a set of recipes or techniques for combining inputs to produce output. Only efficient techniques qualify for inclusion in the function, however, namely those yielding maximum output from any given combination of inputs.

Production functions apply at the level of the individual firm and the macro economy at large. At the micro level, economists use production functions to generate cost functions and input demand schedules for the firm. The famous profit-maximizing conditions of optimal factor hire derive from such micro-economic functions. At the level of the macro economy, analysts use aggregate production functions to explain the determination of factor income shares and to specify the relative contributions of technological progress and expansion of factor supplies to economic growth.

---

■ For valuable comments on earlier drafts of this article, the author is indebted to his Richmond Fed colleagues Bob Hetzel, Ned Prescott, Pierre-Daniel Sarte, and Alex Wolman. The views expressed herein are the author's and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

The foregoing applications are well known. Not so well known, however, is the early history of the concept. Textbooks and survey articles largely ignore an extensive body of eighteenth and nineteenth century work on production functions. Instead, they typically start with the famous two-factor Cobb-Douglas version

$$P = bL^k C^{1-k}.$$

That version dates from 1927 when University of Chicago economist Paul Douglas, on a sabbatical at Amherst, asked mathematics professor Charles W. Cobb to suggest an equation describing the relationship among the time series on manufacturing output, labor input, and capital input that Douglas had assembled for the period 1889–1922.<sup>1</sup>

The resulting equation

$$P = bL^k C^{1-k}$$

exhibited constant returns to scale, assumed unchanged technology, and omitted land and raw material inputs. With its exponents  $k$  and  $1 - k$  summing to one, the function seemed to embody the entire marginal productivity theory of distribution. The exponents constitute the output elasticities with respect to labor and capital. These elasticities, in competitive equilibrium where inputs are paid their marginal products, represent factor income shares that just add up to unity and so exhaust the national product as the theory contends.

The function also seemed to resolve the puzzling empirical constancy of the relative shares. How could those shares remain unchanged in the face of secular changes in the labor force and the capital stock? The function supplied an answer. Increases in the quantity of one factor drive down its marginal productivity and hence its real price. That price falls in the same proportion as the increase in quantity so that the factor's income share stays constant. The resulting share terms  $k$  and  $1 - k$  are fixed and independent of the variables  $P$ ,  $L$ , and  $C$ . It follows that even massive changes in those variables and their ratios would leave the shares unchanged.

From Cobb-Douglas, textbooks and surveys then proceed to the more exotic CES, or constant elasticity of substitution, function

$$P = [kL^{-m} + (1 - k)C^{-m}]^{-1/m}.$$

They observe that the CES function includes Cobb-Douglas as a special case when the elasticity, or flexibility, with which capital can be substituted for labor or vice versa approaches unity.

---

<sup>1</sup> Even before Douglas's collaboration with Cobb, his research assistant at Chicago, Sidney Wilcox, had devised in 1926 the formula  $P = [L^2 + C^2]^{1/2\epsilon} L^k C^h$ , where the exponents  $\epsilon$ ,  $k$ , and  $h$  sum to unity. Wilcox's function reduces to the Cobb-Douglas function in the special case when  $\epsilon$  is zero, but not otherwise (see Samuelson 1979, p. 927).

Finally, the texts arrive at functions that allow for technological change. The simplest of these is the Tinbergen-Solow equation. It prefixes a residual term  $e^{rt}$  to the simple Cobb-Douglas function to obtain

$$P = e^{rt}L^kC^{1-k}.$$

This term captures the contribution of exogenous technological progress, occurring at trend rate  $r$  over time  $t$ , to economic growth. Should new inventions and innovations fail to materialize exogenously like manna from heaven, however, more complex functions are available to handle endogenous technical change. Of these and other post-Cobb-Douglas developments, texts and surveys have much to say. Of the history of production functions before Cobb-Douglas, however, they are largely silent.

The result is to foster the impression among the unwary that algebraic production functions are a twentieth century invention. Nothing, however, could be further from the truth. On the contrary, the idea, if not the actuality, of such functions dates back at least to 1767 when the French physiocrat A. R. J. Turgot implicitly described total product schedules possessing positive first partial derivatives, positive and then negative second partial derivatives, and positive cross-partial derivatives. Thirty years later, Parson Thomas Malthus presented his famous arithmetic and geometric ratios (1798), which imply a logarithmic production function. Likewise, a quadratic production function underlies the numerical examples that David Ricardo (1817) used to explain the trend of the relative shares as the economy approaches the classical stationary state. In roughly the same period, pioneer marginalist Johann Heinrich von Thünen hypothesized geometrical series of declining marginal products implying an exponential production function. Before he died in 1850, Thünen wrote an equation expressing output per worker as a function of capital per worker. When rearranged, his equation yields the Cobb-Douglas function.

Others besides Thünen presaged modern work. In 1877 a mathematician named Hermann Amstein derived from a production function the first-order conditions of optimal factor hire. Moreover, he employed the Lagrangian multiplier technique in his derivation. And in 1882 Alfred Marshall embedded an aggregate production function in a prototypal neoclassical growth model. From the mid-1890s to the early 1900s a host of economists including Philip Wicksteed, Léon Walras, Enrico Barone, and Knut Wicksell used production functions to demonstrate that the sum of factor payments distributed according to marginal productivity exactly exhausts the total product. One of these writers, A. W. Flux, introduced economists to Leonhard Euler's mathematical theorem on homogeneous functions. Finally, exemplifying the adage that no scientific innovation is christened for its true originator, Knut Wicksell presented the Cobb-Douglas function at least 27 years before Cobb and Douglas presented it.

The following paragraphs trace this evolution and identify specific contributions to it. Besides exhuming lost or forgotten ideas, such an exercise

may serve as a partial corrective to the tendency of textbooks and surveys to neglect the early history of the concept. One thing is certain. Algebraic production functions developed hand-in-hand with the theory of marginal productivity. That theory progressed from eighteenth century statements of the law of diminishing returns to late nineteenth and early twentieth century proofs of the product-exhaustion theorem.

Each stage saw production functions applied with increasing sophistication. First came the idea of marginal productivity schedules as derivatives of a production function. Next came numerical marginal schedules whose integrals constitute particular functional forms indispensable in determining factor prices and relative shares. Third appeared the pathbreaking initial statement of the function in symbolic form. The fourth stage saw a mathematical production function employed in an aggregate neoclassical growth model. The fifth stage witnessed the flourishing of microeconomic production functions in derivations of the marginal conditions of optimal factor hire. Sixth came the demonstration that product exhaustion under marginal productivity requires production functions to exhibit constant returns to scale at the point of competitive equilibrium. Last came the proof that functions of the type later made famous by Cobb-Douglas satisfy this very requirement. In short, macro and micro production functions and their appurtenant concepts—marginal productivity, relative shares, first-order conditions of factor hire, product exhaustion, homogeneity and the like—already were well advanced when Cobb and Douglas arrived.

## 1. PRODUCTION FUNCTIONS IMPLICIT IN VERBAL STATEMENTS OF THE LAW OF DIMINISHING RETURNS

The notion of an algebraic production function is implicit in the earliest verbal statements of the operation of the law of diminishing returns in agriculture. A. R. J. Turgot, the French physiocratic economist who served as Louis XVI's Minister of Finance, Trade, and Public Works for a year until dismissed for enacting free-market reforms against the wishes of the king, provided the best of these early statements. In his 1767 *Observations on a Paper by Saint-Péray*, Turgot discusses how variations in factor proportions affect marginal productivities.<sup>2</sup>

Suppose, he writes, that equal increments of the variable factor capital are applied to a fixed amount of land. Each successive increment adds a positive increase to output such that capital's marginal productivity is positive. But that marginal productivity, which at first rises with increases in the capital-to-land

---

<sup>2</sup> On Turgot's discovery of the law of diminishing returns, see Lloyd (1969, p. 22), Niehans (1990, pp. 75–76), and Schumpeter (1954, pp. 259–61). On the production function implicit in Turgot's discovery, see Schumpeter (1954, p. 1036).

ratio, eventually attains a peak and then falls until it reaches zero. At that latter point, the total product of capital—the sum of the marginal products—is at a maximum.

Here is the first clear articulation of the law of variable proportions, or diminishing marginal productivity. Although Turgot applied the law strictly to capital, he realized that it holds for any variable factor including labor. He also recognized a corollary proposition, namely that increases in any factor raise the marginal productivities of the other cooperating factors, which now have more of the first factor to work with. Thus additions to capital, while eventually lowering capital's own marginal productivity, raise the marginal productivities of labor and land.

### **Turgot's Production Function**

Marginal productivity, when expressed mathematically, is the first-order partial derivative of the production function with respect to the input in question, or

$$P/C.$$

And the rate of change of that marginal productivity, again with respect to the associated input, is the second-order partial derivative

$$[P/C]/C = {}^2P/C^2.$$

Finally, the response of an input's marginal productivity to changes in complementary inputs is a cross-partial derivative

$$[P/C]/L = {}^2P/CL.$$

From what has been said above, it follows that Turgot implicitly described a production function possessing positive first partial derivatives, positive then negative second partial derivatives, and positive cross-partial derivatives. His function, with its initially rising marginal productivity of capital, differs from Cobb-Douglas. In Cobb-Douglas, of course, the marginal productivity of a variable factor declines monotonically from the outset so that the second partial derivative is always negative. Also, Turgot's function, because of the fixity of land, cannot exhibit constant returns to scale like Cobb-Douglas.

## **2. PRODUCTION FUNCTIONS IMPLICIT IN NUMERICAL TABLES AND SERIES**

More than 30 years after Turgot, English classical economists independently rediscovered his notion of production functions obeying the law of variable proportions. Unlike him, however, they expressed the concept numerically. Thus several British classicals, though presenting no explicit mathematical production functions, nevertheless used hypothetical numerical examples and series

that imply specific functional forms. A logarithmic function underlies Thomas Malthus's famous arithmetical and geometrical series, which he used to illustrate the law of diminishing returns. In his 1798 *An Essay on the Principle of Population*, Malthus wrote that population, if unchecked, tends to increase indefinitely over time at the geometric ratio 1, 2, 4, 8, 16, 32, 64, 128, 256, 512 . . . . Food output, on the other hand, increases at the arithmetic ratio 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 . . . .

### Thomas Malthus's Logarithmic Production Function

Let  $L$  denote the labor force or its proxy, the population. Similarly, let  $P$  denote food output and  $t$  denote time, normalized so that one unit is the interval required for population to double. (Malthus estimated this doubling time to be 25 years.) Then the equation

$$L = 2^t$$

generates Malthus's geometric series for population as time  $t$  assumes successive values of 0, 1, 2, 3, etc. Similarly, his arithmetic series for food evolves from the equation

$$P = t + 1.$$

Treating the labor force  $L$  and total product  $P$  in the spirit of Malthus as interdependent, interacting variables, one can reduce the two equations to a single logarithmic expression.<sup>3</sup> Solve the second equation for time  $t$ , substitute the result into the first equation, take logarithms, and then solve for  $P$  to obtain the production function

$$P = f(L) = 1 + (1/\log 2)\log L = 1 + (\text{constant})\log L.$$

This production function establishes no absolute upper limit to output. But it does display continuously falling marginal and average productivities of labor. These productivities,

$$dP/dL = f'(L) = (1/\log 2)(1/L)$$

and

$$P/L = f(L)/L = (1/L) + (1/\log 2)(\log L/L),$$

respectively, approach zero asymptotically as the labor force becomes very large. Here are Malthusian diminishing returns with a vengeance.

Malthus put his diminishing-returns production function to immediate use. He employed it to rationalize his minimum-subsistence theory of wages. He

---

<sup>3</sup> George Stigler ([1952] 1965, p. 193) was the first to call attention to Malthus's production function. Peter Lloyd (1969, pp. 22–26) presents an expanded treatment.

argued that labor-force size responds to gaps between actual and subsistence wages. Its response keeps wages at subsistence. Thus above-subsistence wage rates act to raise birth rates, lower death rates, and spur labor-force growth. Because of diminishing returns, however, the extra workers reduce labor's marginal productivity and hence the real wage rate to subsistence. Conversely, below-subsistence wages lead to starvation, low birth rates, and labor-force decline. Fewer workers mean higher marginal productivity of labor, thereby restoring wages to subsistence.

Other classical economists seized on Malthus's population mechanism. Thus was born the classical notion of an unlimited, or infinitely elastic, long-run supply of labor at the subsistence wage rate.

### David Ricardo

Malthus was hardly the only classical economist to work with production functions exhibiting diminishing returns. David Ricardo was the most prominent of the many others who did so. His famous theory of growth and distribution in the economy's progress toward the stationary state rests on a quadratic production function yielding linearly declining marginal and average product schedules. Thus, in his 1817 *Principles of Political Economy and Taxation*, he combines his particular production function with Malthus's minimum-subsistence wage theory to predict that scarcity of land ultimately will bring growth to a halt.

According to Ricardo, growth ceases when diminishing returns to capital applied to scarce land lower capital's real reward to a minimum consistent with zero net investment. At this point, the incentive to invest as well as the means to finance investment vanish and the economy approaches the classical stationary state.

In constructing his production function  $P = f(L)$ , Ricardo assumed that labor and capital combine in rigidly fixed proportions. Each worker, for example, comes equipped with a shovel. The resulting composite input labor-and-capital  $L$  then combines with uniformly fertile land in variable proportions to generate diminishing returns. Ricardo believed that diminishing returns in agriculture were powerful. Indeed, he thought they were so powerful as to overwhelm increasing returns in manufacturing stemming from technological progress and the division and specialization of labor. For that reason, he concentrated on the agricultural sector and omitted variables representing technological progress from his production function. Fixed land, too, was omitted on the grounds that it was a constant rather than a variable. Finally, Ricardo drew no distinction between the aggregate production function for the whole economy and the corresponding micro function for the representative farm. He simply viewed the aggregate function as a scaled-up version of the micro function and treated the economy as one giant farm.

***Ricardo's Quadratic Production Function***

Like Malthus, Ricardo presents his function in the form of a numerical example rather than an algebraic equation.<sup>4</sup> His *Principles* displays a table showing hypothetical marginal products of successive homogeneous doses of labor-and-capital  $L$  applied to land of uniform fertility. The first dose produces 180 units of output. Each succeeding dose contributes 10 fewer units than its immediate predecessor—the second dose contributing 170 units, the third 160 units, and so on. These numbers imply the linearly declining marginal productivity schedule

$$dP/dL = f'(L) = 190 - 10L,$$

which, upon integration, yields the quadratic production function

$$P = f(L) = 190L - 5L^2.$$

One property is absolutely crucial to Ricardo's theory of the trend of relative shares as the economy approaches the stationary state. The function's associated average product schedule

$$P/L = f(L)/L = 190 - 5L$$

declines at half the rate of the marginal product schedule so that the ratio of marginal to average product falls as  $L$  increases.

***Ratio of Marginal to Average Products and the Trend of Relative Shares***

This property, together with Ricardo's assumption that Malthusian population growth keeps the wage rate at subsistence, determines the trend of relative shares in his model. For it is easy to show that the shares going to rent on the one hand and wages plus profit on the other vary inversely and directly, respectively, with the ratio of the marginal to the average product of labor-and-capital. After all, land's absolute real rental income  $R$  is simply what remains of total product  $P = f(L)$  after the variable composite factor  $L$  receives its marginal product  $f'(L)$ . That is,

$$R = f(L) - Lf'(L).$$

Dividing through by total product gives rent's relative share

$$R/P = [f(L) - Lf'(L)]/f(L) = 1 - \{f'(L)/[f(L)/L]\}$$

as one minus the term in braces. This latter term represents the combined share of total product going to labor and capital together. It is nothing other than the crucial ratio of the marginal to the average product of the composite variable

---

<sup>4</sup> Haim Barkai deduced this function from Ricardo's tables in 1959. Both he and Blaug (1985, pp. 88–92, 103–05, 118–21) discuss how Ricardo used it to predict the trend of rent's distributive share as the economy approaches the stationary state.



$L$ . Since the ratio falls with increasing applications of  $L$ , it follows that rent's share rises while the combined share of wages and profit falls.

Of the combined share, the wage component must rise and the profit component fall. The reason is simple. Since the Malthusian mechanism holds the wage rate at subsistence, the total wage bill consisting of the wage rate times the work force grows proportionally with the number of workers. Output, however, grows less than proportionally to labor because of diminishing returns. As a result, the ratio of the wage bill to total product, namely labor's share, increases with  $L$ . And with the relative shares of rent and wages both rising, it follows that the remaining relative share of profit necessarily falls.

### *Approach to the Classical Stationary State*

Eventually, profit and its relative share fall to a minimum, perhaps zero. There both the means and the incentive to finance net investment vanish. At that point, profit is just sufficient to maintain rather than to increase the capital stock.

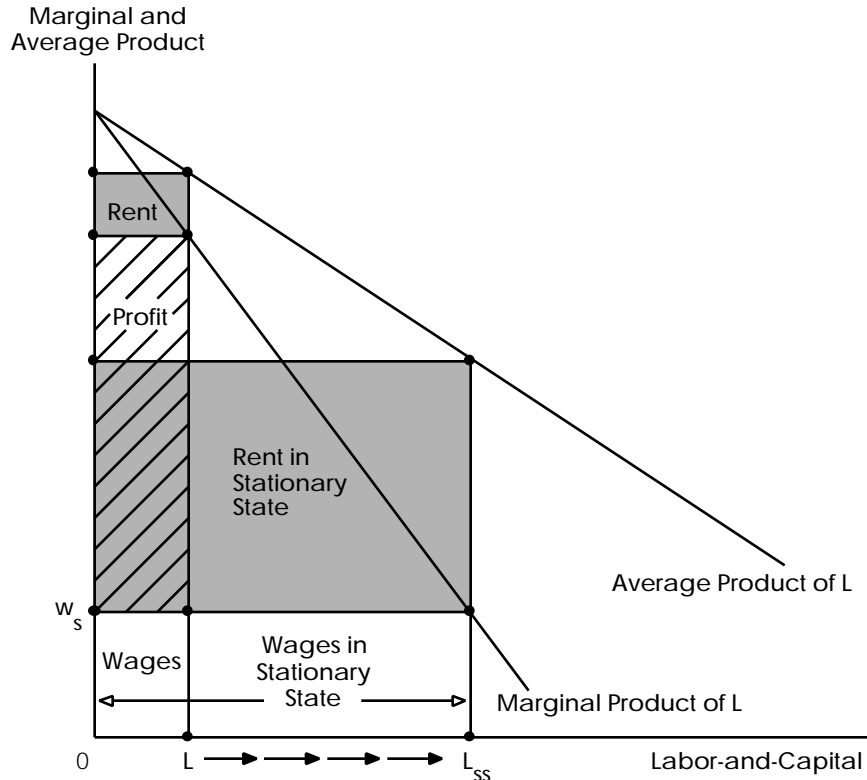
With capital formation stymied, growth halts. Here is Ricardo's prediction that diminishing returns stemming from scarcity of land overwhelm increasing returns due to technological progress and so lead inevitably to the classical stationary state. His pessimistic, dismal prophecy derives directly from a production function exhibiting linearly declining marginal and average returns (see Figure 1).

### *Ricardian Rent Theory as Incomplete Marginalism*

As a marginal productivity theory of factor pricing and distribution, Ricardo's rent analysis left much to be desired. True, it did establish the marginal principle. It stressed that capitalist farmers should cultivate land to the point where the incremental return to the last dose of labor-and-capital applied just equals the cost of the dose. But it employed marginal analysis only to determine the joint payment going to labor and capital combined.

To account for the rewards going to each factor separately, Ricardo had to resort to other explanations. He relied on the Malthusian minimum-subsistence theory to determine labor's wage rate and income share. Similarly, he explained capital's profit rate as a pure residual, namely what remained of the marginal product of labor-and-capital after deduction of subsistence wages. Likewise, he viewed land's rental rate as a surplus determined by the gap between the average and marginal products of the variable factor, or alternatively, by the superior productivity of the intramarginal doses of the factor.

In short, Ricardo resorted to subsistence and residual theories to determine factor prices. Marginal productivity served only to split the total product into its rent and non-rent components. What was needed was someone to transform the primitive, incomplete marginalism of classical Ricardian theory into comprehensive neoclassical marginal productivity.

**Figure 1 Ricardo's Theory of the Trend of the Relative Shares**

Ricardo's quadratic production function generates linearly declining marginal and average product schedules. Total product consists of the rectangular area inscribed under the average product curve at any given level of labor-and-capital. The composite labor-and-capital input receives its marginal product, of which labor gets subsistence wages  $w_s L$  and capital gets the cross-hatched area as profit. Land rent accounts for the remaining rectangle bounded by the gap between the marginal and average product curves. Such rent consists of the surplus product of the inframarginal doses of the composite factor. Bidding for scarce land transfers the surplus to landowners.

Initially, with labor-and-capital at  $L$ , profit exists to fuel capital formation and growth. Growth stops when the composite factor expands to its stationary-state level  $L_{ss}$ . There diminishing returns reduce profit to zero, thus extinguishing the source of growth. In the classical stationary state, all product accrues to labor and land. The widening gap between the marginal and average product curves ensures that the profit share falls while the rent and wage shares rise as the economy expands along the horizontal axis.

### **Johann Heinrich von Thünen's Contributions to Marginal Productivity Theory**

Credit for doing so goes to the German mathematical economist, location theorist, and agronomist Johann Heinrich von Thünen (1783–1850), whose work on production functions was far ahead of its time. A true original, he owed nothing to the productivity doctrines of Malthus and Ricardo. Having read neither writer, Thünen constructed his theories fresh from the detailed records that he kept for his own agricultural estate. In a lifelong effort to identify empirically the exact relations of production on his farm, he applied marginal analysis to all factor inputs and prices. The entire neoclassical theory of production and distribution traces its origin to Volume II of his great book *The Isolated State*.

As the earliest neoclassical marginalist, Thünen boasts several distinctions. He was the first to apply the differential calculus to productivity theory and perhaps the first to use it to solve economic optimization problems.<sup>5</sup> He was likewise the first to interpret marginal productivities essentially as partial derivatives of the production function. In so doing, he made explicit what was merely implicit in Turgot's analysis.

Decomposing Ricardo's composite input into its separate components, Thünen was the first to treat labor and capital symmetrically, to show that each is subject to diminishing incremental returns, and to state that labor's marginal productivity is an increasing function of the quantity of capital per worker. Moreover, he was the first to state precisely that capital's real reward and labor's too are determined by the additional product resulting from the last increment of each factor hired, all other factors held constant. Likewise, he was the first to show that when capital's interest rate is determined marginally, wages may appear to be a residual. Conversely, when the wage rate is determined marginally, interest appears as a residual. Unlike Ricardo, who assumed fixed factor proportions, Thünen stressed that labor and capital are substitutes for each other and can be combined in variable proportions.

Always Thünen insisted that economic efficiency requires that factors be hired until the ratio of their respective marginal products equals the ratio of their unit prices. Always he held that net revenue peaks when each input's marginal value product matches its marginal factor cost. In short, he relentlessly applied the principle that equimarginal resource allocation maximizes the total product.

---

<sup>5</sup> The vexed question of priority of discovery raises its head here. Thünen evidently used the differential calculus to solve an optimization problem as early as 1824. But in that same year Thomas Perronet Thompson employed rudimentary elements of the calculus to compute the optimal inflation tax. Even earlier, in a book published in 1815, Georg von Buquoy used the calculus to determine the optimal plowing depth of the soil. In any case, Thünen's optimization calculus remained unpublished until 1850, by which time Augustin Cournot's 1838 *Researches into the Mathematical Principles of the Theory of Wealth* had eclipsed it. On these contributions see Niehans (1990, pp. 173–74) and the sources cited there.

These contributions identify him as the true founder of neoclassical marginal productivity theory.<sup>6</sup>

### Thünen's Exponential Production Functions

Bolstering his theory with numbers derived from agricultural experiments on Tellow, his estate in Mecklenburg, Thünen presented tables depicting the marginal productivities of labor  $L$ , capital  $C$ , and fertilizer  $F$  applied to fixed land. His tables show the marginal productivities declining at constant geometrical rates. In other words, his experiments suggested to him that successive unit increases of any variable input, the others held constant, add to output a constant fraction of the amount added by the preceding unit. For labor the fixed fraction was two-thirds, for capital nine-tenths, and for fertilizer one-half.

Let  $r$  denote the fractional ratio between successive marginal products of any variable factor. Then Thünen's schedule of the factor's incremental returns constitutes the terms of the decreasing geometric series  $a, ar, ar^2, ar^3, \dots, ar^{n-1}$ . Here  $a$  is the marginal product of the first unit of the factor,  $ar$  the marginal product of the second,  $ar^2$  that of the third and so on until we reach the last, or  $n$ th, unit whose marginal product is  $ar^{n-1}$ .

Using the sum-of-the-series formula

$$S = [a(1 - r^n)/(1 - r)]$$

to sum the  $n$  marginal products gives the factor's total product schedule as

$$P = A(1 - r^n),$$

where the letter  $A$  denotes the constant term  $a/(1 - r)$  and the exponent  $n$  is the number of units of the variable factor hired. Since  $r$  is a fraction such that  $r^n$  becomes zero when  $n$  becomes infinite, it follows that  $A$  is the limit that the sum  $A(1 - r^n)$  approaches as the number of factor units  $n$  becomes indefinitely large. The upshot is that the factor's total product asymptotically converges to the finite maximum  $A$ .

Exactly the same analysis holds for each and every variable input. Consequently, when all factors—labor, capital, and fertilizer—are allowed to vary simultaneously, the production function underlying Thünen's numerical examples can be expressed as

$$P = f(L, C, F) = A(1 - 2/3^L)(1 - 9/10^C)(1 - 1/2^F),$$

where the exponents denote the amounts of each factor employed.

---

<sup>6</sup> On Thünen and his contributions to marginal productivity theory, see Blaug (1985, pp. 322–24, 449, 616–17), Dickinson (1969), Leigh (1968), Niehans (1990, pp. 164–75), and Schumpeter (1954, pp. 465–68).

The foregoing applies when inputs vary in units of discrete size. Should those units be infinitely divisible and so continuously variable, then the term  $e^{-k}$  replaces the fraction  $r$  in the production function. Here  $k$  denotes the instantaneous rate of decline of marginal productivity and  $e^{-k}$  is the factor of proportionality over the unit interval. The result is that each input's total product schedule becomes

$$P = A(1 - e^{-kn}),$$

and the corresponding Thünen production function is

$$P = f(L, C, F) = A(1 - e^{-.405L})(1 - e^{-.105C})(1 - e^{-.693F}).$$

This function, like its discrete counterpart, possesses two properties. Output is zero when any factor is zero. Output approaches its maximum level  $A$  as all factors are increased indefinitely.<sup>7</sup>

### Rediscovery of Thünen's Exponential Functions

Thünen's exponential production functions together with their associated marginal schedules passed largely unnoticed for more than 40 years after their publication in 1863. They (but not their authorship) were rediscovered by two agricultural scientists, W. J. Spillman and E. A. Mitscherlich, working independently in the early twentieth century. Spillman labeled Thünen's decreasing geometric series the "law of the diminishing increment" and wrote out the corresponding marginal and total product equations for fertilizer and irrigation water. Mitscherlich christened the same phenomenon the "law of the soil." This law he represented by the equation

$$dP/dF = k(A - P),$$

expressing fertilizer's marginal productivity  $dP/dF$  as a constant fraction of the (diminishing) gap between maximum  $A$  and actual current levels  $P$  of output. Upon integration, his equation yields Thünen's total product schedule

$$P = f(F) = A(1 - e^{-kF})$$

for the continuous case. Neither author, however, was aware of Thünen's earlier formulation of these concepts.

### 3. THE FIRST ALGEBRAIC PRODUCTION FUNCTION

In addition to the functions implicit in his numerical examples, Thünen wrote down the first explicit algebraic production function to appear in print (see Lloyd [1969], pp. 31–33). As presented in Chapter 2 of the second part of

<sup>7</sup> Lloyd (1969, pp. 26–31) provides a complete account of Thünen's exponential production functions and their derivation. See also Stigler (1946, p. 125) for a textbook treatment.

Volume II of *The Isolated State*, the function evidently replaces the geometric series of marginal products he presented in the first part of that same volume. Expressed in per-worker form, his function is

$$p = hq^n,$$

where  $p$  is output per worker,  $h$  is a constant parameter determined by such considerations as the fertility of the soil and the strength and diligence of the workers who till it,  $q$  is capital per worker (the capital-to-labor ratio), and the exponent  $n$  is a fraction between zero and one.

It turns out that Thünen's production function is none other than the Cobb-Douglas function  $P = bL^k C^{1-k}$  in disguise. For, when one multiplies both sides of Thünen's equation by labor  $L$ , one obtains

$$P = pL = hLq^n = hL(C/L)^n = hL^{1-n}C^n.$$

The resulting function

$$P = hL^{1-n}C^n$$

is virtually the same as the Cobb-Douglas function. The conclusion is inescapable. Credit for presenting the first Cobb-Douglas function, albeit in disguised or indirect form, must go to Thünen in the late 1840s rather than to Douglas and Cobb in 1928. All of which goes to show that there is nothing new under the sun. Or as statistician Stephen M. Stigler expressed it in his famous Law of Eponymy, "No scientific discovery is named for its original discoverer."

Thünen's equation states that production requires inputs of both labor and capital such that labor working alone produces nothing. Thünen was uncomfortable with this result. Surely labor unaided by capital has some productivity, however low. To ensure that labor's output is positive even if capital is zero, Thünen modified his production function to read

$$p = h(1 + q)^n,$$

or, when multiplied through by labor  $L$ ,

$$P = hL^{1-n}(L + C)^n.$$

This equation, which Thünen estimated empirically for his own agricultural estate and which he declares he discovered only after more than 20 years of fruitless search, states that labor produces something even when unequipped with capital.<sup>8</sup>

---

<sup>8</sup> The variable  $q$  in this equation expresses the capital-labor ratio when capital is measured in units of work effort. Alternatively, Thünen uses the letter  $k$  to denote the ratio when capital is expressed in the workers' means of subsistence. In this latter case, his formula becomes

$$p = h(g + k)^n,$$

where  $g$  is a positive constant not necessarily equal to unity.

#### 4. FIRST USE OF AN AGGREGATE PRODUCTION FUNCTION IN A NEOCLASSICAL GROWTH MODEL

For all its brilliance and originality, Thünen's productivity analysis had little impact on his contemporaries and immediate successors. Some were intimidated by its formidable mathematics and shunned it for that reason. Thünen's own countrymen largely ignored it because it was theoretical and thus ran counter to the anti-theoretical bias of the dominant German Historical School. Still others overlooked it because it was hidden amidst the profusion of cryptically written notes, comments, digressions, repetitions, numerical examples, and mathematical formulas that constituted the disjointed and cumbersome narrative of *The Isolated State*. Another reason for neglect was that Thünen's readers concentrated on his celebrated but misguided formula  $w = \sqrt{ap}$  to the exclusion of his other work. That formula, which Thünen thought sufficiently important to have engraved on his tombstone, specified the natural wage  $w$  as the geometric mean of the worker's minimum subsistence  $a$  and his average product  $p$ . Preoccupied with the formula, readers tended to overlook Thünen's genuine contributions to production theory.

One economist who was influenced, however, was Alfred Marshall. He acknowledges his debt to Thünen in a fragment preserved in the 1925 *Memorials of Alfred Marshall*. There he credits Augustin Cournot with teaching him pure analytic technique and Thünen with teaching him economics. Confessing that he derived more of his ideas from Thünen than from Cournot, he declares that he reveres Thünen above all his masters.

##### Marshall's Growth Model

Given his indebtedness to Thünen's productivity ideas, it is hardly surprising to find Marshall, in a note written in 1881 or 1882, employing an aggregate production function. His function appears in what is best described as a prototypical neoclassical growth model. In that model, Marshall expresses aggregate annual output or real national product  $P$  as a function of four determinants. These are, respectively, the number  $L$  times the average efficiency  $E$  of the working population (the work force measured in efficiency units), the accumulated stock of capital  $C$ , the level of technology or state of the arts of production  $A$ , and the fertility of the soil  $F$ , which Marshall treats as a fixed constant. In symbols,

$$P = f(L \cdot E, C, A, F).$$

Taken together, the growth rates of the arguments of Marshall's function determine the growth rate of aggregate output. Marshall expresses these input growth rates as time derivatives— $dL/dt$ ,  $dE/dt$ ,  $dC/dt$ ,  $dA/dt$ —each treated as a function of several relevant variables including wage and interest rates, the standard of comfort, time, and the arguments in the production function itself. In principle, the resulting dynamic system could be solved to yield secular growth paths for population, capital, technology, and output.

Marshall of course did not solve the system or investigate the qualitative properties of its growth paths. His formulation was too sketchy for such exercises. Nevertheless, he did incorporate an aggregate production function in what may be regarded as the first neoclassical growth model. And he did so at least 60 years before Tinbergen and Solow, generally regarded as the fathers of the neoclassical growth model. Unfortunately, however, Marshall never published his model and thus denied himself the credit he deserved. It remained for J. K. Whitaker to discover the model among the unpublished manuscript notes deposited in the Marshall Library at Cambridge University and to publish it in 1975.

## 5. PRODUCTION FUNCTIONS USED TO DERIVE THE CONDITIONS OF OPTIMAL FACTOR HIRE

Ironically, Marshall penned his model at the very time when the focus of economic theory had shifted from aggregate growth to individual optimization and allocation. Thus it was not macro but rather micro production functions that began to appear with increasing frequency in the economics literature of the 1890s. Paving the way was the so-called marginal revolution of the early 1870s. That event saw the marginalist triumvirate of William Stanley Jevons, Carl Menger, and Léon Walras apply what in essence was the calculus of constrained optimization to the consumer's utility function. The result was the derivation of the marginal utility theory of consumer's demand. Second-generation marginalists soon realized that those same optimization techniques might be applied to the production function of the individual firm to find the profit-maximizing or cost-minimizing conditions of factor hire. Thünen's marginal productivity theory was born again.

### Hermann Amstein

Even before the 1890s, however, a University of Lausanne mathematician named Hermann Amstein had worked out virtually the entire theory of marginal productivity in modern algebraic dress. He did so in response to a request from his colleague Léon Walras, who sought Amstein's help in formulating mathematically the least-cost conditions of factor hire. In a letter of January 6, 1877, Amstein responded with a near-perfect analysis, complete with partial derivatives and Lagrangian multipliers, of cost minimization subject to a production function constraint.

His analysis went as follows. Define unit cost of production  $U$  as the ratio of total cost to output. Total cost consists of the sum of the factor inputs each multiplied by its unit price. Let the wage rate  $w$  and the interest rate  $i$  denote the unit prices of labor and the services of capital, respectively. Competitive firms of course take these and all factor prices as given. Then Amstein argued that the problem is to find the input quantities  $L, C, \dots$  that minimize unit cost



$$U = (wL + iC + \dots)/P$$

for any given level of production  $\bar{P}$ .

Today economists solve this problem in three steps. First they take the cost function  $U$ . Then they subtract from it the production function less the given level of output  $f(L, C, \dots) - \bar{P}$ , all multiplied by an arbitrary Lagrangian multiplier  $\lambda$ . Finally, they minimize the resulting Lagrangian expression

$$Z = U - \lambda[f(L, C, \dots) - \bar{P}]$$

by setting its first partial derivatives equal to zero.

That is precisely what Amstein did, with one exception. He suppressed the given product term  $\bar{P}$  while setting the production function  $f$  at naught. Then he minimized the resulting Lagrangian

$$Z = (wL + iC + \dots) - \lambda f(L, C, \dots)$$

by setting its first partial derivatives equal to zero. This operation yielded him the first order conditions

$$w - \lambda (f/L) = 0 \text{ and } i - \lambda (f/C) = 0, \dots$$

These conditions, when rewritten as

$$w = \lambda (f/L) \text{ and } i = \lambda (f/C), \dots$$

and divided by each other so as to cancel out the  $\lambda$ s,

$$w/i = (f/L)/(f/C), \dots$$

identify the least-cost combination of factor inputs as that which equates the ratio of factor prices with the ratio of factor marginal productivities or, alternatively, which renders the marginal product per last dollar spent on each factor

$$(f/L)/w = (f/C)/i = \dots$$

the same across factors.

Amstein's work is a milestone in the history of production functions. Here was the first use of the Lagrangian multiplier technique in economics.<sup>9</sup> Here also was the first rigorous derivation of the least-cost conditions of factor hire from a constrained cost function.

---

<sup>9</sup>It was not the first to appear in print, however. John Creedy (1980) reports that Francis Edgeworth, in his 1877 *New and Old Methods of Ethics*, employed the Lagrangian multiplier technique to find the distribution of income that maximizes aggregate community satisfaction or utility. And, in his 1881 *Mathematical Psychics*, Edgeworth again used the technique to derive the contract-curve solution of Pareto-efficient allocations according to which each trader maximizes his utility subject to the condition that the other trader's utility remains constant.

These innovations, however, went completely unnoticed. For Walras, who at the time had hardly progressed beyond elementary analytical geometry and was just beginning to teach himself the rudiments of calculus, knew too little mathematics to understand Amstein's formulation and to take advantage of it. And Amstein himself knew too little economics to appreciate the significance of his demonstration and to prepare it for publication. For these reasons his contribution remained unknown until William Jaffé discovered it in the Lausanne archives and published it in 1964. The result was to delay the progress of production theory for at least 12 years. Not until 1889 was a production function employed again in an optimization problem. And not until the 1920s were Lagrange multipliers seen again in production-function analysis. Technique here ran ahead of its potential users until they became convinced of the gain from mastering it.

### Francis Y. Edgeworth and Profit Maximization

Amstein derived the conditions of optimal factor hire from the competitive firm's constrained cost function. He solved a cost-minimization problem in which the production function entered as a constraint. By contrast, the next group of writers derived the factor-hire conditions from the firm's profit function. They solved a profit-maximization problem in which the production function entered as a component of gross revenue. Profit, or net revenue, they defined as the difference between gross revenue and total cost. Gross revenue consisted of product price multiplied by output as represented by the production function. Total cost consisted of the sum of factor inputs each multiplied by its factor price.

Thus Francis Edgeworth, in his 1889 *Journal of the Royal Statistical Society* article "On the Application of Mathematics to Political Economy," stated that the entrepreneur acts to maximize the profit or net revenue expression

$$f(C, T) - iC - rT.$$

Here  $f$  is gross revenue, or output valued at its given competitive market price (implicitly assigned a value of unity by Edgeworth),  $C$  is capital,  $T$  is land,  $i$  is the interest rate or price of the services of capital, and  $r$  is the rent-per-acre price of land. Maximizing this expression by setting its first partial derivatives equal to zero, Edgeworth obtained the conditions

$$f/C = i \text{ and } f/T = r.$$

In short, profit maximization requires hiring factors up to the point where they just pay for themselves, namely where their marginal value products equal their prices.

### Arthur Berry and William Ernest Johnson

Cambridge lecturers Arthur Berry, a mathematician, and William Ernest Johnson, a logician, philosopher, and economic theorist, then extended Edgeworth's analysis in four ways. First, they increased the number of inputs in the production function. Thus Berry's function, which appears in his 1891 paper "The Pure Theory of Distribution," contains separate symbols for capital as well as for labor and land, both subdivided into unlimited kinds and qualities. Likewise, Johnson's production function, as presented in his 1891 piece entitled "Exchange and Distribution," embodies a potentially unlimited number of variable factor inputs  $V_i$ .

Second, Berry and Johnson incorporated product price into the entrepreneur's profit expression, thus making explicit what Edgeworth had left implicit. Johnson's model is typical of Berry's as well. Let  $\pi$  stand for product price,  $P( )$  for product quantity (the production function),  $w$  for input price,  $V$  for input quantity, and the subscript  $i = 1, 2, 3 \dots$  for the separate inputs. Then the entrepreneur seeks to maximize the profit expression

$$\pi P(V_i) - \sum w_i V_i \quad (i = 1, 2, 3, \dots),$$

where the first term is gross revenue and the second is total cost. Maximization yields the first-order conditions

$$\pi(P/V_1) = w_1, \quad \pi(P/V_2) = w_2, \quad \pi(P/V_3) = w_3, \text{ etc.}$$

Together, these state that each factor should be hired to the point where its marginal value product equals its price.

Third, suppose the firm operates under imperfect, rather than perfect, competition. Facing a downward-sloping demand curve, the firm finds the selling price of its product now varies inversely with output. Correspondingly, its marginal revenue now always lies below its price. Berry and Johnson noted that this special case necessitates replacing product price with marginal revenue in the first-order factor-hire conditions. Those conditions then read: hire factors up to the point where their marginal revenue products, or marginal physical products multiplied by marginal revenue, equal their factor prices.

Fourth, Berry and Johnson indicated how the factor-hire conditions might be incorporated into simple general equilibrium systems complete with commodity demand functions, labor supply functions, and full-employment conditions. These models no doubt influenced Edgeworth. For, in his 1894 review of Friedrich von Wieser's book *Natural Value*, he incorporated dual production functions into a two-good, two-factor model of general equilibrium. Doing so, he showed that the equality of marginal value product per last dollar spent on each factor must be the same across all goods as well as factors.

Taken together, these contributions constitute what Joseph A. Schumpeter (1954, p. 1032n) termed "a considerable achievement." They show that the

production function already was becoming an essential component of micro models of the business firm by the first half of the 1890s.

## 6. PRODUCTION FUNCTIONS AND THE ADDING-UP PROBLEM

Production functions continued to prove their worth in the latter half of the 1890s when marginalists employed them to resolve the famous adding-up problem of product exhaustion.<sup>10</sup> At stake was nothing less than the logical consistency of the marginal productivity theory of distribution. Would wages, rent, and interest, if each input is paid its marginal product, just add up to and so exhaust the total product as the theory claimed? That is, would the total product exactly be disposed of without residue or shortage?

A positive answer would confirm the consistency of the theory. But a negative answer would refute it. For if the sum of the payments according to marginal productivity exceeded the total product, the excess would go unrealized since no firm could afford to pay out more than is produced. Some inputs would be forced to accept less than their marginal products, contrary to the theory. Conversely, if less was paid out under marginal productivity than was produced, the remaining surplus would have to be distributed on grounds other than marginal productivity, contrary to the theory. Small wonder that marginalists were eager to prove the answer was yes.

### Product Exhaustion under Constant Returns to Scale: Philip H. Wicksteed

First to do so expressly was Philip H. Wicksteed.<sup>11</sup> In his 1894 *An Essay on the Co-ordination of the Laws of Distribution*, Wicksteed proved that product exhaustion holds, provided perfect competition prevails and production functions are linear homogeneous and so exhibit constant returns to scale. Competition ensures that inputs receive their marginal products. And linear homogeneity ensures that the resulting distributive shares sum to the total product.

Had he realized it, Wicksteed could have deduced adding-up directly from Leonhard Euler's famous mathematical theorem on homogeneous functions. That theorem says that any linear homogeneous function can be written as the sum of its first partial derivatives each multiplied by the associated independent

---

<sup>10</sup> George Stigler (1941, Chapter XII) is the standard source on the history of the product exhaustion problem. See also Steedman's (1987) useful treatment.

<sup>11</sup> Already Knut Wicksell, in his 1893 *Value, Capital, and Rent*, had constructed a marginal productivity model that implied a proof of product exhaustion (see Stigler [1941], pp. 289–95). But he failed to make the proof explicit and was content to see Wicksteed receive credit for its discovery in the following year.

variable. In other words, the production function  $P = f(L, C, \dots)$ , if linear homogeneous as Wicksteed thought, can be written as

$$P = (f/L)L + (f/C)C + \dots,$$

where the terms on the right-hand side are factor incomes determined by marginal productivity. Here at once is the proof, ready-made, that Wicksteed sought. Curiously enough, however, he never used it. Owing to his lack of formal mathematical training, he apparently was unaware of the theorem and so made no mention of it. Instead, he sought to prove adding-up, or product exhaustion, by reconciling marginal productivity with Ricardo's theory of intensive rent.

### *Wicksteed's Proof of Product Exhaustion*

Such reconciliation was necessary. For in Ricardo's theory there is no adding-up problem to solve. Rent, as we have seen, is a pure residual in the Ricardian model. It is what is left of the total product after the other (composite) factor, labor-and-capital, has received its marginal product. And with rent determined residually, it is tautologically true that the sum of factor incomes must just add up to the total product. The residual would always adjust to make it so. To transform Ricardo's theory into one in which the adding-up theorem applied, Wicksteed had to prove that Ricardian residual rent was the same as rent as marginal product. This proof would then imply that remuneration of all factors according to their marginal productivities exhausts the total product.

His demonstration required four steps. First, he did what Ricardo had failed to do. He entered land explicitly into the production function by writing the function in per-acre form. That is, he expressed product per acre  $P/T$  as an increasing function of labor-and-capital per acre  $L/T$ . In symbols, he posited

$$P/T = f(L/T)$$

or

$$P = Tf(L/T).$$

Second, he expressed Ricardian rent  $R$  as the residual part of total product remaining after each unit of labor-and-capital  $L$  receives its marginal product payment  $P/L$ . That is,

$$\begin{aligned} R &= P - (P/L)L \\ &= Tf(L/T) - T[f(L/T)/L]L \\ &= Tf(L/T) - T\{f'(L/T)[(L/T)/L]\}L \\ &= Tf(L/T) - T[f'(L/T)(1/T)]L \\ &= Tf(L/T) - f'(L/T)L. \end{aligned}$$

Here is rent income expressed as Ricardian residual.

Third, he expressed land's income alternatively as marginal product. That is, he computed the partial derivative

$$P/T = [Tf(L/T)]/T$$

to represent the marginal product of the last acre in use and multiplied it by the number of acres cultivated  $T$ . The result was the expression

$$[f(L/T) + Tf'(L/T)(-L/T^2)]T,$$

which reduces to

$$Tf(L/T) - Lf'(L/T),$$

precisely the same expression as residual rent. Here is his proof that Ricardian residual rent equals rent as marginal product.

Fourth, to this computed marginal productivity payment to land he adds the corresponding marginal productivity payment to labor-and-capital. He gets

$$Lf'(L/T) + Tf(L/T) - Lf'(L/T),$$

which equals

$$Tf(L/T)$$

or total product  $P$ . Here is his proof that product exhaustion occurs when both factors are paid their marginal products.

#### **A. W. Flux and Euler's Theorem**

After Wicksteed came A.W. Flux. His innovation was to accomplish what Wicksteed had failed to do, namely to deduce the adding-up proposition directly from Euler's theorem. His review of Wicksteed's *Co-ordination*, published in the June 1894 issue of the *Economic Journal*, is absolutely clear on this point.

Let the production function be linear homogeneous such that a scalar increase in all inputs yields the same scalar increase in output. Then, wrote Flux, "Euler's equation gives us at once the result" that output equals the sum of the inputs each multiplied by its partial derivative, or marginal productivity. Put differently, Euler's theorem gives us the result that factor income shares determined by marginal productivity must sum to unity and so absorb the product. Here is the first application of Euler's theorem to production function analysis. Contrary to common belief, it was Flux and not Wicksteed who introduced this theorem to economists.

#### **The Critics: Barone, Edgeworth, Pareto, and Walras**

The Wicksteed-Flux proof of product exhaustion received an inhospitable reception. Critics including Enrico Barone, Francis Edgeworth, Vilfredo Pareto, and Léon Walras attacked its homogeneity assumption. They argued that linear

homogeneity renders adding-up a trivial outcome that holds at every point on the production function regardless of the proportions in which the factors are combined. In other words, homogeneity proves too much and is thus too good to be true. Edgeworth's ([1904], 1925, p. 31) caustic remark was devastating: "There is a magnificence in this generalization which recalls the youth of philosophy. Justice is a perfect cube, said the ancient sage; and rational conduct is a homogeneous function, adds the modern savant."

The critics further noted that the homogeneity proposition yields horizontal long-run marginal and unit cost curves. Such curves render firm size indeterminate. They thus cast doubt on the large-numbers property of competition. Why? Because competitive firms possessing horizontal cost curves can minimize cost at any scale of operation. With no cost advantage to being small or disadvantage to being large, a firm could be of any size. But such firm-size indeterminacy in turn implies indeterminacy of the number of firms in the industry. Conceivably, a few firms might be so large as to monopolize the market, contrary to the assumptions of the competitive model.

Vilfredo Pareto (1897) adduced three additional reasons working against linear homogeneity. First, some factors are in fixed supply. They cannot expand equiproportionally with the others as homogeneity implies. One can, for example, replicate all the elements of a restaurant on the Champs Elysées except the location itself. Second, some inputs come in units that are large and indivisible. Such lumpy inputs can hardly be scaled up or down in proportion to output as homogeneity assumes. An example is a train tunnel that can accommodate a quadrupling of the traffic but that cannot be subdivided into smaller tunnels of the same efficiency to handle a fraction of the traffic. Third, some factors are in a fixed technological relation with the product (iron and iron ore) or with each other (trucks and truck drivers). Their lack of independent variation thwarts the freedom of factor substitution that homogeneity assumes. To Pareto, these reasons were enough to render production functions nonhomogeneous so that they exhibit increasing or decreasing returns to scale.

### **Knut Wicksell's Reconciliation of Nonhomogeneity with Adding-Up**

Knut Wicksell clarified, refined, and considerably amplified the foregoing observations. In so doing, he reconciled product exhaustion with nonhomogeneity. Unlike Wicksteed, who saw constant- and nonconstant-returns production functions as mutually exclusive phenomena, Wicksell (1901, 1902) argued that a firm's production function might exhibit successive stages of increasing, constant, and decreasing returns to scale. These stages correspond to the falling, constant, and rising segments of the firm's U-shaped long-run average cost curve. Free entry of rivals in pursuit of profit forces the competitive firm to operate at the minimum point of this curve. Or what is the same thing, competition forces the firm to operate at the zero-profit point, where its production function

is tangent to a linear homogeneous plane. Here, constant returns and therefore adding-up prevail. In short, product exhaustion is an equilibrium condition that holds at the single point where the firm's nonhomogeneous production function behaves as if it were linear homogeneous.

Wicksell noted, however, that nonhomogeneous production functions for firms are perfectly compatible with a linear homogeneous function for the entire industry. Suppose industry output expands and contracts through the entry and exit of identical firms, each operating at the same minimum unit cost. The result is to trace out a horizontal long-run industry supply curve that looks like it came from a constant-returns production function.

Here was Wicksell's contribution. At one stroke, he solved three problems. He reconciled adding-up with nonhomogeneity of production functions at the level of the individual firm. He then reconciled those functions with homogeneous functions for the industry. In so doing, he justified the use of aggregate linear homogeneous functions such as the Cobb-Douglas function. Finally, he reconciled competitive equilibrium with determinate firm size.

### Product Exhaustion under Nonconstant Returns

The preceding considerations led Barone, Walras, and Wicksell to formulate an alternative proof of product exhaustion. Dispensing with Wicksteed's assumption of linear homogeneity, they instead posited nonhomogeneous production functions yielding U-shaped long-run unit cost curves. They interpreted product exhaustion as an outcome of competitive equilibrium in which firms operate at the minimum point of these curves and charge a price equal to the minimum unit cost.

Their proof, already anticipated by Amstein in 1877, is straightforward. Into the competitive firm's unit cost equation

$$U = (wL + iC + \dots)/P$$

substitute the cost-minimizing conditions of optimal factor hire. Since these conditions state that factor prices equal factor marginal physical products times product price  $\pi$ , such substitution yields the expression

$$U = \pi[(f/L)L + (f/C)C + \dots]/P.$$

Divide both sides by  $U$  and multiply both sides by  $P$  to obtain

$$P = (\pi/U)[(f/L)L + (f/C)C + \dots].$$

Note that free entry in long-run competitive equilibrium dictates that firms produce at the minimum, or zero profit, point on their unit cost functions where product price  $\pi$  equals unit cost  $U$ . The upshot is that the term  $\pi/U$  equals one and the equation reduces to the product-exhaustion condition

$$P = (f/L)L + (f/C)C + \dots$$



Wicksell was right. Evidently, competitive equilibrium ensures that even nonhomogeneous production functions deliver product exhaustion with factor shares adding up to unity. It is enough that the functions be momentarily homogeneous at the equilibrium point.<sup>12</sup>

## 7. WICKSELL'S ANTICIPATION OF THE COBB-DOUGLAS FUNCTION

The adding-up controversy had at least one important unintended consequence. It advanced knowledge of production functions to the point where the Cobb-Douglas equation, heretofore known only to Thünen, was within easy grasp of serious scholars. Wicksell is the key figure here. It was he who essentially transformed Thünen's implicit or disguised version of the Cobb-Douglas function into its exact or final form. And he did so on at least five occasions, the first appearing 27 years before and the last appearing four years before Cobb-Douglas. Owing to him, economists hardly had to wait for the equation to appear in 1928. Instead, they could refer to Wicksell, who was already using it at the turn of the century.

It is easy to trace the evolution of the equation in Wicksell's writings (see Olsson [1971], Sandelin [1976], and Velupillai [1973]). Like Thünen before him, Wicksell begins, in his 1896 *Finanztheoretische Untersuchungen*, by

<sup>12</sup> John R. Hicks (1932, pp. 234–39) proved as much without referring to product price. Minimize unit cost

$$U = (wL + iC \dots) / P$$

by setting its first partial derivatives at zero. Make use of the definition that the sum of factor prices times factor quantities equals total cost or  $UP$ . The resulting partial derivatives

$$U/L = (1/P^2)[Pw - UP(P/L)] = (1/P)[w - U(P/L)]$$

and

$$U/C = (1/P^2)[Pi - UP(P/C)] = (1/P)[i - U(P/C)]$$

when set at zero reduce to

$$w = U(P/L)$$

and

$$i = U(P/C).$$

Substitute these into the unit cost equation to get

$$U = U[L(P/L) + C(P/C) + \dots] / P.$$

Multiply both sides by  $P$  and divide both sides by  $U$  to get the product-exhaustion expression

$$P = (P/L)L + (P/C)C + \dots$$

presenting a per-worker version of the function.<sup>13</sup> Four years later, in his 1900 *Ekonomisk Tidskrift* piece on “Marginal Productivity as the Basis for Distribution in Economics,” he advances to an exact replica of the function. He notes that the Wicksteed product-exhaustion formula for labor and land

$$P = (P/L)L + (P/T)T$$

is a partial differential equation that has the general solution

$$P = Lf(T/L).$$

He then cites as one example of this class of functions the Cobb-Douglas equation

$$P = L^\alpha T^\beta,$$

where the exponents  $\alpha$  and  $\beta$  sum to unity.

The Cobb-Douglas formula reappears in Volume 1 of his 1901 *Lectures on Political Economy*. There he adds that if the exponents  $\alpha$  and  $\beta$  together exceed or fall short of unity, the factor shares will respectively over- and under-exhaust the product. He then insists that competition, by forcing firms to operate at minimum unit cost where constant returns prevail, ensures that the exponents sum to one as required by product exhaustion. Similarly, in correspondence with his colleague David Davidson in 1902, he uses the Cobb-Douglas function to prove that, with constant returns to scale, the joint marginal product of labor, land, and capital together equals the sum of their separate marginal products (see Uhr [1991]).

Again, in his 1916 article “The ‘Critical Point’ in the Law of Decreasing Agricultural Productivity,” he employs the Cobb-Douglas function

$$P = L^\alpha T^\beta C^\gamma$$

to reconcile constant returns to scale with diminishing returns to proportions. You have constant returns to scale when a 10 percent increase in all inputs increases product by

$$10(\alpha + \beta + \gamma) = 10(1) = 10 \text{ percent.}$$

---

<sup>13</sup> His function is

$$p = ch^m t^k b^v,$$

where  $p$  is output per worker,  $c$  is a constant,  $h$  is the land-to-labor ratio,  $t$  and  $b$  are the lengths of the investment periods of labor and land, respectively, and the exponents  $m$ ,  $k$ , and  $v$  are fractions. Multiplying both sides by labor  $L$  yields

$$pL = P = cL(T/L)^m t^k b^v = cL^{1-m} T^m t^k b^v.$$

This function is Cobb-Douglas in labor and land but not in the investment periods. Nor should it be since such periods are hardly factor inputs logically parallel to labor and land.

You have diminishing returns to proportions when a 10 percent increase in both labor and capital, land held constant, increases product by

$$10(\alpha + \gamma) < 10 \text{ percent.}$$

Since output increases by less than the 10 percent increase in labor and capital, it follows that the average product of those inputs decreases. It does so because each unit of augmented labor and capital must work with a smaller amount of cooperating land.

Finally, in his 1923 review of Gustaf Akerman's doctoral dissertation *Realkapital und Kapitalzins*, Wicksell writes the Cobb-Douglas function as

$$P = cL^\alpha C^\beta,$$

with the exponents  $\alpha + \beta$  adding up to one. Clearly, if priority of discovery were the criterion, the names of Wicksell and Thünen should precede those of Cobb and Douglas when attached to the function. Credit should go to Cobb and Douglas not for inventing the function itself but for showing that it provides a good description of the aggregate data.

## 8. SOME FINAL OBSERVATIONS

The preceding discussion has concentrated exclusively on major landmarks in the history of production functions before Cobb-Douglas. In so doing, it undoubtedly has overlooked other milestones.

For example, nothing was said about the fixed-proportion production functions of Richard Cantillon (1755), Léon Walras (1874), and Gustav Cassel (1918). These functions foreshadowed modern Leontief functions and share the same features. Production is characterized by rigidly fixed input coefficients that rule out factor substitution. Such coefficients specify input requirements per unit of output. Thus if one hour of labor  $L$  (assisted of course with the required amounts of cooperating inputs) can produce ten units of product  $P$  such that each unit of product requires a tenth of an hour of labor, then  $L/P = 1/10$  is the input coefficient of labor in producing output. Similar coefficients hold for other required inputs.

Denote the production coefficients of labor, capital, and land as  $l$ ,  $c$ , and  $t$ , respectively. Then Cantillon, Walras, Cassel, and Leontief could write the production function as

$$P = \min (L/l, C/c, T/t).$$

Here the terms  $L/l$ ,  $C/c$ , and  $T/t$  are, respectively, the largest outputs producible from the quantities of labor, capital, and land available. The smallest of these terms determines the level of output. Why? Because with fixed factor proportions, output is limited by the relatively scarcest factor, just as the size of a cake is limited by the recipe ingredient in shortest supply. Given the quantity

of the limitational ingredient, extra units of the other ingredients would not increase output; their marginal products are zero. At the point of limitation, output absorbs inputs precisely in the ratio  $l : c : t$ .

By 1900, however, most economists, including Walras himself, were balking at the evidently unrealistic notion of fixed proportions and zero factor substitution. Already they were employing variable-coefficient functions rather than fixed-coefficient ones. Fixed production coefficients, however, made a comeback in the 1950s and 1960s in linear programming and input-output models.

The preceding paragraphs also failed to mention the British physicist Lord Kelvin's 1882 engineering production function for the transmission of electric power (see Smith [1968], p. 515). Kelvin's pioneering work (see Appendix) foreshadowed modern engineering production functions for gas and heat transmission, steam power production, metal cutting, and batch reactor chemical processes.

Nevertheless, enough has been said to document the main contention of the article, namely that algebraic production functions long predate Cobb-Douglas. At least 18 economists from seven countries over a span of 160 years either presented or described such functions before Cobb-Douglas. Seen in this perspective, the Cobb-Douglas function and its more recent successors represent the culmination of a long tradition rather than the beginning of a new one.

---



---

## APPENDIX

Lord Kelvin (William Thompson) related electric power output  $P$  (the quantity of electric energy delivered) to two factors, namely power input  $I$  and the size  $S$  of the copper cable, or conductor, through which the electric current is transmitted. Power output is that part of power input not lost through frictional heating of the cable. Such loss varies directly with the square of power output and inversely with the size (weight or volume) of the cable. The result is the implicit production function

$$P = I - (kP^2/S),$$

where  $k$  is a constant that depends on the length and resistance properties of the cable. When solved explicitly for output,

$$P = (S/2k)\{[1 + (4kI/S)]^{1/2} - 1\},$$

Kelvin's function ascribes quantity of electric power delivered to three determinants, namely power input, cable size or weight, and the constant of resistance. From this function derives Kelvin's famous law stating that the conductor cable reaches its optimum size when the annual interest cost of the copper invested in the cable equals the value of the energy lost annually through heating of the cable.

---



---

**REFERENCES**

- Amstein, Hermann. Correspondence on Marginal Productivity Theory, 1877, in William Jaffé, “New Light on an Old Quarrel,” *Cahiers Vilfredo Pareto*, No. 3, 1964, pp. 94–97, and William Jaffé, ed., *Correspondence of Léon Walras and Related Papers, 1857–1883*, Vol. I. Amsterdam: North-Holland, 1965, pp. 516–20, reprinted in William J. Baumol and Stephen M. Goldfeld, eds., *Precursors in Mathematical Economics: An Anthology*. London: The London School of Economics and Political Science, 1968, pp. 309–12.
- Barkai, H. “Ricardo on Factor Prices and Income Distribution in a Growing Economy,” *Economica*, vol. 26 (August 1959), pp. 240–50.
- Barone, Enrico. “Sur un livre récent de Wicksteed,” 1896, French translation by Léon Walras in *Cahiers Vilfredo Pareto*, No. 3, 1964, pp. 68–73.
- Berry, Arthur. “The Pure Theory of Distribution,” *Report of the Sixtieth Meeting of the British Association for the Advancement of Science*, 1891, reprinted in William J. Baumol and Stephen M. Goldfeld, eds., *Precursors in Mathematical Economics: An Anthology*. London: The London School of Economics and Political Science, 1968, pp. 314–15.
- Blaug, Mark. *Economic Theory in Retrospect*, 4th ed. Cambridge: Cambridge University Press, 1985.
- Cantillon, Richard. *Essai sur la nature du commerce en général*, 1755, edited with English translation and other material by Henry Higgs. London: Macmillan (for the Royal Economic Society), 1931.
- Cassel, Gustav. *Theoretische Sozialökonomie*. Leipzig: C. F. Winter, 1918, translated into English as *Theory of Social Economy*. London: T. F. Unwin, 1923, new revised edition London: E. Benn, 1932.
- Cobb, Charles W., and Paul H. Douglas. “A Theory of Production,” *American Economic Review*, vol. 18 (Supplement) (March 1928), pp. 139–65.
- Cournot, Antoine Augustin. *Researches into the Mathematical Principles of the Theory of Wealth*, 1838, translated by Nathaniel T. Bacon. New York: Macmillan, 1929, reprinted, New York: Augustus Kelley, 1971.
- Creedy, John. “The Early Use of Lagrange Multipliers in Economics,” *Economic Journal*, vol. 90 (June 1980), pp. 371–76.
- Dickinson, H. D. “Von Thünen’s Economics,” *Economic Journal*, vol. 79 (December 1969), pp. 894–902.

- Edgeworth, Francis Y. "The Theory of Distribution," *Quarterly Journal of Economics*, 1904, reprinted in Francis Y. Edgeworth, *Papers Related to Political Economy*, Vol. 1. London: Macmillan (for the Royal Economic Society), 1925.
- \_\_\_\_\_. Review of *Natural Value* by Friedrich von Wieser, *Economic Journal*, 1894, reprinted in Francis Y. Edgeworth, *Papers Related to Political Economy*, Vol. 3. London: Macmillan (for the Royal Economic Society), 1925.
- \_\_\_\_\_. "On the Application of Mathematics to Political Economy," *Journal of the Royal Statistical Society*, vol. 52 (December 1889), pp. 538–76, reprinted in Francis Y. Edgeworth, *Papers Related to Political Economy*, Vol. 2, London: Macmillan (for the Royal Economic Society), 1925.
- \_\_\_\_\_. *Mathematical Psychics*. London: C. Kegan Paul & Co., 1881.
- \_\_\_\_\_. *New and Old Methods of Ethics*. Oxford: James Parker & Co., 1877.
- Flux, A. W. Review of Philip H. Wicksteed, *Essay on the Co-ordination of the Laws of Distribution*, in *Economic Journal*, vol. 4 (June 1894), pp. 308–13, reprinted in William J. Baumol and Stephen M. Goldfeld, eds., *Precursors in Mathematical Economics: An Anthology*. London: The London School of Economics and Political Science, 1968, pp. 326–31.
- Hicks, John R. *The Theory of Wages*. New York: Macmillan, 1932.
- Jaffé, William. "New Light on an Old Quarrel," *Cahiers Vilfredo Pareto*, No. 3, 1964, pp. 61–102.
- Johnson, William E. "Exchange and Distribution," *Cambridge Economic Club*, 1891, reprinted with brief commentary in William J. Baumol and Stephen N. Goldfeld, eds., *Precursors in Mathematical Economics: An Anthology*. London: The London School of Economics and Political Science, 1968, pp. 316–20.
- Kelvin, Lord (William Thompson). "On the Economy of Metal in Conductors of Electricity," 1882, Report of 51st meeting, held August-September 1881, British Association for the Advancement of Science, *Reports* 51, pp. 526–28.
- Leigh, Arthur H. "Thünen, Johann Heinrich von," in David L. Sills, ed., *International Encyclopedia of the Social Sciences*, Vol. 16. Macmillan and Free Press, 1968, pp. 17–20.
- Lloyd, Peter J. "Elementary Geometric/Arithmetic Series and Early Production Theory," *Journal of Political Economy*, vol. 77 (January/February 1969), pp. 21–34.
- Malthus, Thomas Robert. *An Essay on the Principle of Population*, 1st ed. London, 1798.

- Mitscherlich, E. A. "Das Gesetz des Minimums und das Gesetz des abnehmenden Bodenertrages," *Landw. Jahrb.*, vol. 38 (1909), pp. 537–52.
- Niehans, Jurg. *A History of Economic Theory*. Baltimore: Johns Hopkins University Press, 1990.
- Olsson, Carl-Axel. "The Cobb-Douglas or the Wicksell Function?" *Economy and History*, vol. 14 (1971), pp. 64–69.
- Pareto, Vilfredo. *Cours d'économie politique*, II. Paris, 1897.
- Pigou, Arthur Cecil, ed. *Memorials of Alfred Marshall*. London: Macmillan, 1925, reprinted, New York: Kelley and Millman, 1956.
- Ricardo, David. *On the Principles of Political Economy and Taxation*. London: Murray, 1817, reprinted in Piero Sraffa, ed., *The Works and Correspondence of David Ricardo*, Vol. 1. Cambridge: Cambridge University Press, 1951.
- Samuelson, Paul A. "Paul Douglas's Measurement of Production Functions and Marginal Productivities," *Journal of Political Economy*, vol. 87 (October 1979), pp. 923–39.
- Sandelin, Bo. "On the Origin of the Cobb-Douglas Production Function," *Economy and History*, vol 19 (No. 2, 1976), pp. 117–23.
- Schumpeter, Joseph A. *History of Economic Analysis*. London: Allen & Unwin, 1954.
- Smith, Vernon L. "Production," in David L. Sills, ed., *International Encyclopedia of the Social Sciences*, Vol. 12. Macmillan and Free Press, 1968, pp. 511–18.
- Spillman, W. J., ed. *The Law of Diminishing Returns*. Yonkers-on-Hudson, N.Y.: World Book Co., 1924.
- Steedman, Ian. "Adding-up Problem," in John Eatwell, Murray Milgate, and Peter Newman, eds., *The New Palgrave: A Dictionary of Economics*, Vol. 1. London: Macmillan, 1987, pp. 21–22.
- Stigler, George J. "The Ricardian Theory of Value and Distribution," *Journal of Political Economy*, vol. 60 (June 1952), pp. 187–207, reprinted in George J. Stigler, *Essays in the History of Economics*. Chicago: University of Chicago Press, 1965, pp. 156–97.
- \_\_\_\_\_. *The Theory of Price*. New York: Macmillan, 1946.
- \_\_\_\_\_. *Production and Distribution Theories: The Formative Period*. New York: Macmillan, 1941.
- Stigler, Stephen M. "Stigler's Law of Eponymy," *Transactions of the New York Academy of Sciences*, 2d series, vol. 39 (April 1980), pp. 147–58.

- Thünen, Johann Heinrich von. *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*. 3 volumes. Jena, Germany: Fischer, 1930. A partial translation of *The Isolated State in Relation to Agriculture and Political Economy*, Vol. 2, "The Natural Wage and Its Relation to the Rate of Interest and to Economic Rent" is included in Bernard W. Dempsey, *The Frontier Wage: The Economic Organization of Free Agents*. Chicago: Loyola University Press, 1960.
- Turgot, Anne Robert Jacques. *Observations sur le Mémoire de M. de Saint-Péray*, Limoges, 1767, in E. Daire, ed., *Oeuvres de Turgot*, Paris, 1844, pp. 414–33, translated as "Observations on a Paper by Saint-Péray on the Subject of Indirect Taxation," in P. D. Groenewegen, *The Economics of A. R. J. Turgot*, The Hague, 1977, pp. 43–95.
- Uhr, Carl H. "David Davidson: The Transition to Neoclassical Economics," in Bo Sandelin, ed., *The History of Swedish Economic Thought*. London: Routledge, 1991, pp. 44–75.
- Velupillai, Kumaraswamy. "The Cobb-Douglas or the Wicksell Function?—A Comment," *Economy and History*, vol. 16 (1973), pp. 111–13.
- Walras, Léon. "Note on Mr. Wicksteed's Refutation of the English Theory of Rent," Appendix III of *Eléments d'économie politique pure*, 3d ed. Lausanne, 1896.
- \_\_\_\_\_. *Eléments d'économie politique pure*, 1st ed. Lausanne: F. Rouge, 1874.
- Whitaker, J. K., ed. *The Early Economic Writings of Alfred Marshall, 1867–1890*, Vol. 2. New York: The Free Press, 1975.
- Wicksell, Knut. *Über Wert, Kapital, und Rente*, Jena, Germany: G. Fischer, 1893, translated by S. H. Frowein as *Value, Capital, and Rent*. London: Allen & Unwin, 1954, reprinted, New York: Augustus M. Kelley, 1970.
- \_\_\_\_\_. "The 'Critical Point' in the Law of Decreasing Agricultural Productivity," 1916, in E. Lindahl, ed., *Selected Papers on Economic Theory by Knut Wicksell*. London: Allen & Unwin, 1958.
- \_\_\_\_\_. "On the Problem of Distribution," 1902, in E. Lindahl, ed., *Selected Papers on Economic Theory by Knut Wicksell*. London: Allen & Unwin, 1958.
- \_\_\_\_\_. "Marginal Productivity as the Basis for Distribution in Economics," *Ekonomisk Tidskrift*, 1900, in E. Lindahl, ed., *Selected Papers on Economic Theory by Knut Wicksell*. London: Allen & Unwin, 1958, pp. 93–121.
- \_\_\_\_\_. "Real Capital and Interest," 1923, Review of Gustaf Akerman's *Realkapital und Kapitalzins*, in Lionel Robbins, ed., *Lectures on Political Economy*, Vol 1: *General Theory*, Appendix 2. London: Routledge & Kegan Paul, 1934.



———. *Föreläsningar i nationalekonomi*, Vol. I, Lund, 1901, translated by E. Classen in the third Swedish edition, 1928, and in Lionel Robbins, ed., *Lectures on Political Economy*, Vol. 1: *General Theory*. London: Routledge & Kegan Paul, 1934.

———. *Finanztheoretische Untersuchungen nebst Darstellung und Kritik des Steurowesens Schwedens*. Jena, Germany: G. Fischer, 1896.

Wicksteed, Philip H. *An Essay on the Co-ordination of the Laws of Distribution*. London: Macmillan & Co., 1894, revised with an introduction by Ian Steedman, Aldershot, England: Edward Elgar, 1992.