Chapter 1: Introduction

1.1 Background

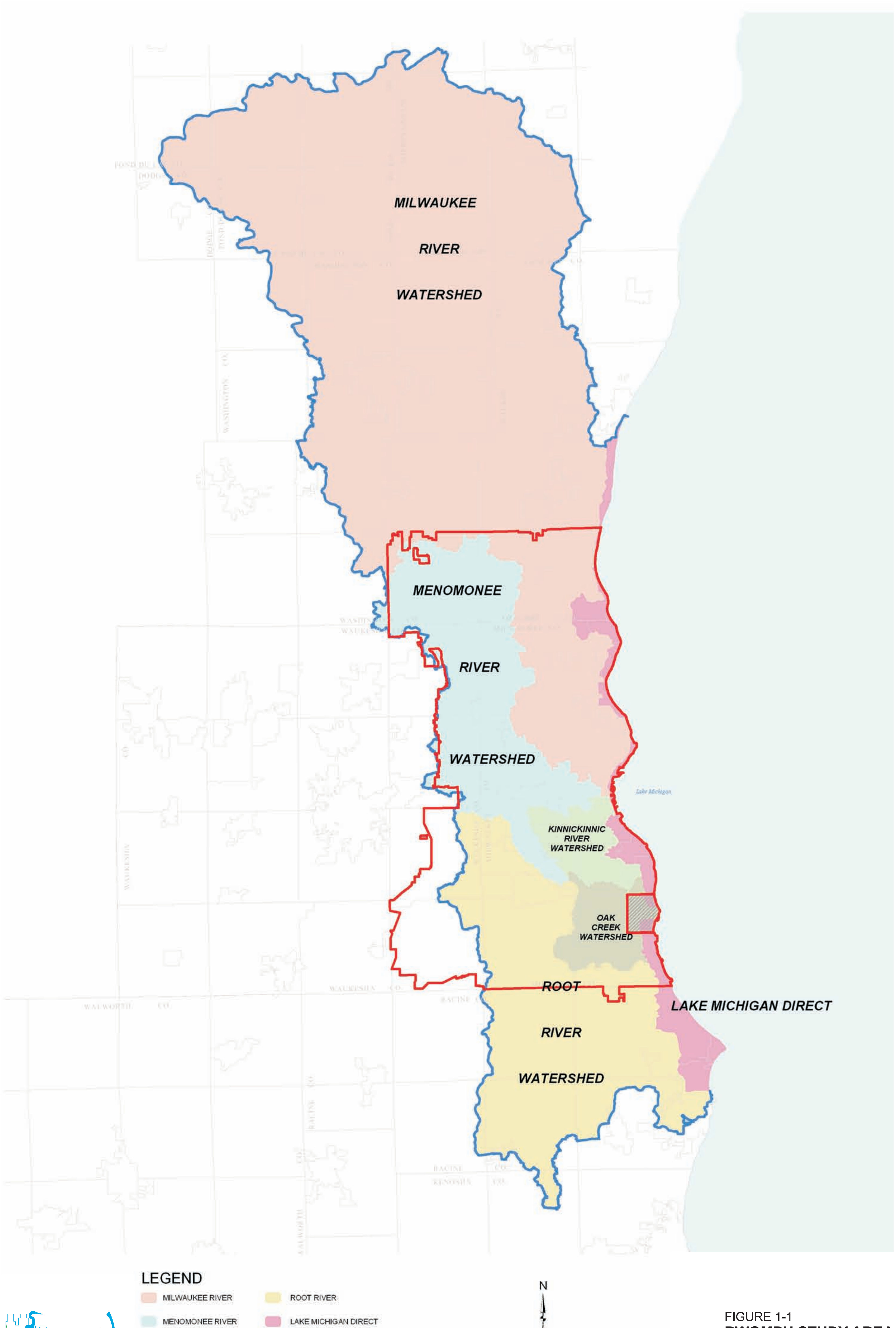
The *State of the Art Report* (SOAR) has been prepared as part of the coordinated efforts by the Southeastern Wisconsin Regional Planning Commission (SEWRPC) and the Milwaukee Metropolitan Sewerage District (MMSD). The two resulting projects are the SEWRPC Regional Water Quality Management Plan Update (RWQMPU) and MMSD 2020 Facilities Plan (2020 FP). This integrated planning effort is called the Water Quality Initiative (WQI), and this SOAR report is a critical element for both planning efforts that form the WQI.

The WQI uses watershed-based planning for the greater Milwaukee watersheds (GMW), which include the Milwaukee River, the Menomonee River, the Kinnickinnic River, the Root River, Oak Creek, the Milwaukee Harbor estuary, and a portion of nearshore Lake Michigan and its direct drainage area. Figure 1-1 shows the RWQMPU study area and 2020 FP planning boundaries.

The Wisconsin Department of Natural Resources (WDNR), MMSD, and SEWRPC cooperatively developed an inclusive, open and science-based approach to carry out the WQI. Under the approach, which was formalized under a February 19, 2003 Memorandum of Understanding (1) among these agencies, a collaborative planning process was used to evaluate the cost effectiveness and feasibility of sets of technologies, i.e., the alternatives, distinct from the responsibility for their implementation, in order to identify an alternative that offers the greatest improvement in the water resource at the least total cost to society. This effort will result in an update to the regional water quality management plan for the GMW and support the 2020 FP.

To determine the most cost effective method to achieve the greatest improvement in surface water quality, the planning programs used the production theory (see Appendix 1A of this chapter) to evaluate water pollution control technologies. This theory was determined to be the most logical way to analyze multiple "production" functions (e.g., water pollution control technologies) with varied inputs, such as point source and nonpoint source water pollution, and outputs, such as reduction of overflow volumes and/or water pollutants.

Although not included in the SOAR analysis, the impacts on water quality as a result of reducing overflows were modeled using the Loading Simulation Program in C++ (LSPC) model, the Estuarine Coastal and Ocean Model-Sediment Transport (ECOMSED), and the Row-Column AESOP (RCA) model. The water quality impacts were evaluated in the screening alternatives and the preliminary alternatives and were considered in the selection of the components for the Recommended Plan. The concentrations of pollutants in sanitary sewer overflows (SSOs) and combined sewer overflows (CSOs) used in the water quality model are shown in Tables 3A-1 through 3A-3 of Appendix 3A, *Point Source Technologies*. Additional information regarding the impacts to water quality as a result of reducing overflows can be found in Chapter 9, *Alternative Analysis* of the *Facilities Plan Report* and Chapter IX, *Alternative Plan Description and Evaluation* of SEWRPC Planning Report No. 50, *A Regional Water Quality Management Plan Update for the Greater Milwaukee Watersheds.*



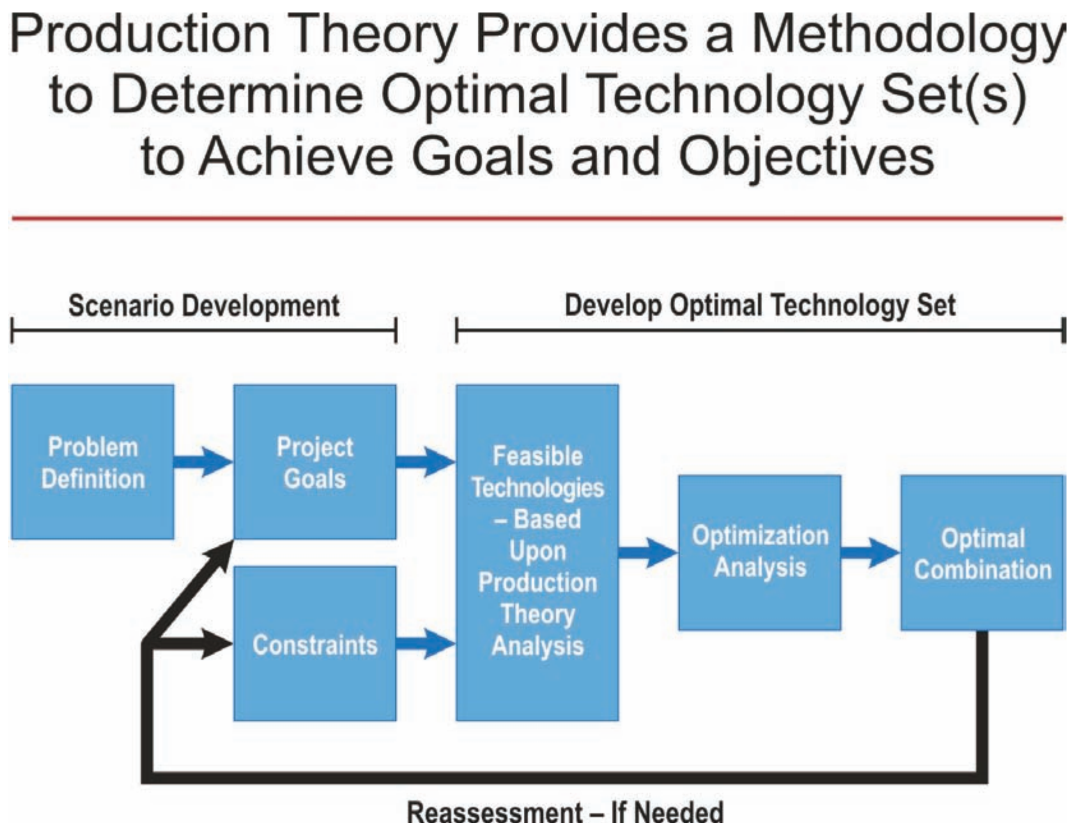
Preserving The Environment • Improving Water Quality

FIGURE 1-1
RWQMPU STUDY AREA AND MMSD 2020 PLANNING BOUNDARY
2020 STATE OF THE ART REPORT
4/25/07 SOAR_01.0001.07.04.25.cdr

Production theory may in fact be the only scientifically-based method that will achieve the desired result of the SOAR analysis, which is to use a science and data-based economic analysis to identify the pollution reduction technologies that will produce the most cost efficient outcomes in terms of water quality improvement.

The production theory is an economic evaluation technique first conceived in the 19th century and used to evaluate the outputs that can be obtained from various amounts and combinations of factor inputs. A detailed discussion of production theory is found in Appendix 1A, *Production Theory*. While the technologies could have been analyzed using scientific and engineering judgment, it would have been difficult to quantify due to the vast number of technologies that were evaluated and the various modeling techniques that were used in the WQI.

In the production theory, each technology is described by a cost function, a production function, a cost benefit relationship, and its interaction with other control technologies. Using this information, technologies that address similar water quality indicators are compared. Also, combinations of technologies working in series or in parallel can be evaluated. Finally, the best combination of technologies that maximizes water quality benefits and minimizes costs is determined. Figure 1-2 shows the general production theory methodology and how it can be applied to meet a project's goals and objectives.



Preserving The Environment • Improving Water Quality

FIGURE 1-2
PRODUCTION THEORY METHODOLOGY
2020 STATE OF THE ART REPORT
4/25/07 SOAR_01.0002.07.04.25.cdr

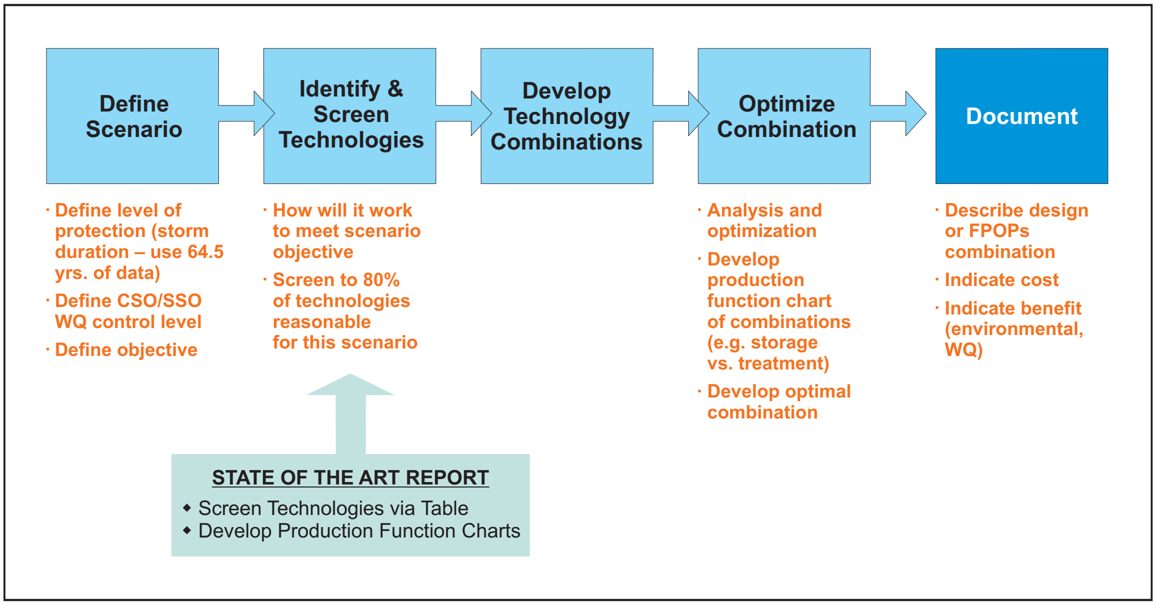
1.2 Objectives of the Production Theory Analysis

The objective of the analysis was to evaluate the potential of both individual technologies and multiple technology sets to improve water quality in the GMW for the WQI. To achieve this objective, approximately 170 technologies were included in the initial evaluation. From these, select technologies were further evaluated using the production theory.

The ultimate objective of the WQI effort is to identify the specific combination of technologies that provide the greatest water quality benefit at the lowest total cost, considering both the technology screening in this report and input from the stakeholders. Combinations of technologies were first developed to analyze extreme "what-if" situations, or "bookends," called screening alternatives (originally called scenarios). These screening alternatives were defined and developed to address questions from the media, public and other stakeholders regarding "what if" situations such as, "if we ended overflows – what would it cost and what would the water quality be?" The screening alternatives were developed as a sensitivity analysis to determine: a) what would happen to water quality if we only concentrated on eliminating overflows or if we only tried to aggressively control nonpoint source pollution and, b) the general costs of each "what-if" situation. The screening alternatives were not intended to be feasible solutions, but rather evaluations of extreme conditions, or "bookends." The screening alternatives were evaluated to develop preliminary alternatives based upon a broader set of considerations, including stakeholder-inspired goals and objectives as well as current regulations. A detailed discussion of the screening alternatives is provided in the *Facilities Plan Report*, Appendix 9A, *Screening Alternatives.* The relationship of SOAR to the development of screening alternatives for both the RWQMPU and the 2020 FP is shown in Figure 1-3. A similar process was used to develop the preliminary alternatives. A detailed discussion of the alternative development process is presented in Chapter IX of SEWRPC Planning Report No. 50, *Alternative Plan Description and Evaluation* and Chapter 9, *Alternative Analysis* of the *Facilities Plan Report*.

This report is intended to explain the technology analysis performed for the WQI. It is not intended to cover the following issues:

- Roles and Responsibilities – this is discussed in Chapter 7, *Goals and Objectives* of the *Facilities Plan Report*.
- Classification of sewer separation technologies (e.g., Level I or Level II) – this is discussed in Appendix 10A, *CSO Long-term Control Plan* of the *Facilities Plan Report*.



CSO = combined sewer overflows
SSO = sanitary sewer overflows
WQ = water quality
FPOPs = facilities, programs, operational improvements and policies

Preserving The Environment • Improving Water Quality

FIGURE 1-3
RELATIONSHIP OF SOAR TO THE RWQMPU AND THE 2020 FP SCREENING ALTERNATIVES

References

(1) *Memorandum of Understanding* between the Milwaukee Metropolitan Sewerage District (District), the Southeastern Wisconsin Regional Planning Commission (SEWRPC), and the Wisconsin Department of Natural Resources (WDNR) for cooperation in the watershed approach to water quality and facilities planning (formalized February 19, 2003)

Appendix 1A:
Production Theory

Appendix 1A: Production Theory

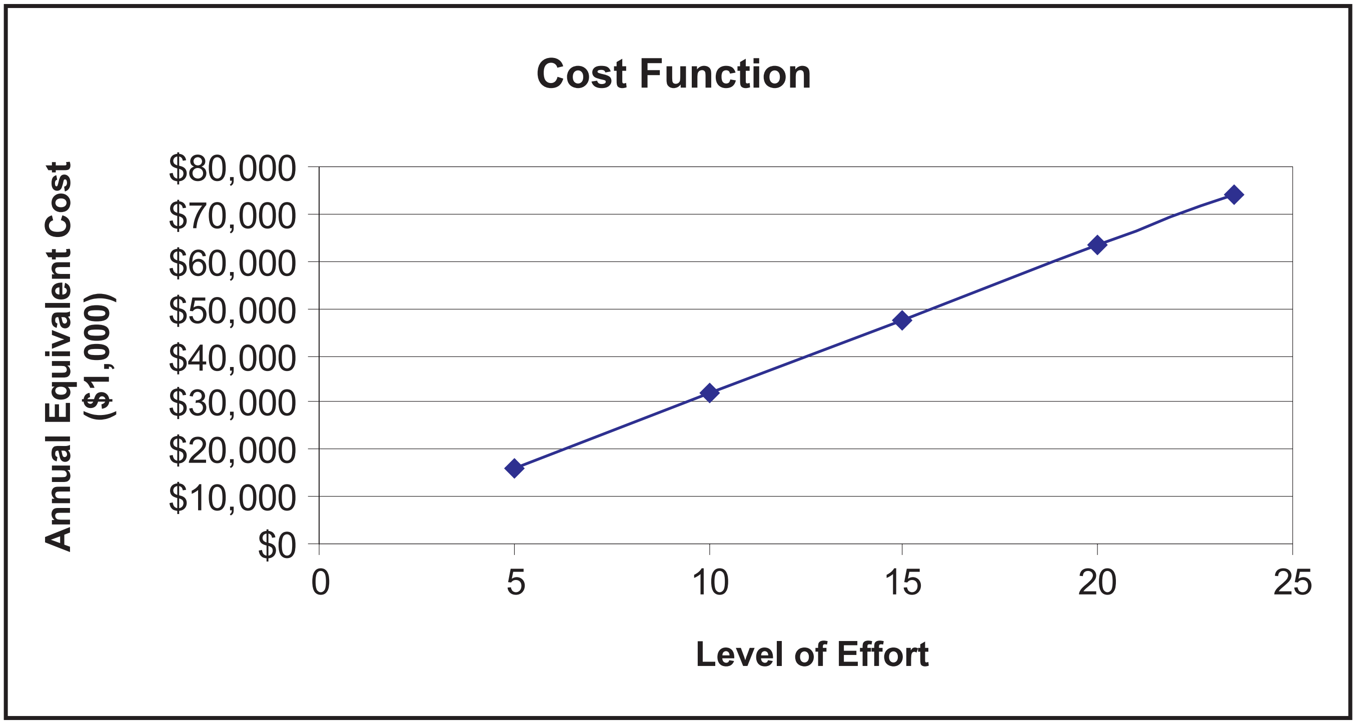
Production theory is a procedure and an economic tool that is used to determine the most efficient way to use resources to achieve a desired goal. The theory is most often taught and used in business applications. For example, it may be used when making decisions involving the cost of producing a product, or when deciding which manufacturing plant of a group of plants is most cost efficient to produce a given product. The basic principle of the production theory is that with variable production techniques a producer can choose various capital-labor combinations, all of which yield the same output but with varying overall costs.

For the purposes of the Water Quality Initiative (WQI) joint planning effort, which resulted in the Regional Water Quality Management Plan Update (RWQMPU) and 2020 Facilities Plan (2020 FP), production theory was used to determine the most efficient combination of technologies to achieve the water quality goals and objectives of the public. Each technology is described by a cost function, a production function, a cost benefit relationship, and its interaction with other control technologies. Using these relationships, an optimal combination of technologies can be selected. A description and examples of these components are provided below.

1A.1 Cost Function

The relationship between the total life cycle cost incurred and the input (level of effort) applied.

Life cycle cost can be expressed in terms of equivalent annual cost or net present worth. The level of effort can be the volume of tunnel storage, square miles of sewer separation or annual lane miles of street sweeping. Figure 1A-1 shows an example of a cost function.



Preserving The Environment • Improving Water Quality

FIGURE 1A-1
EXAMPLE OF COST FUNCTION
2020 STATE OF THE ART REPORT
4/25/07 SOAR_1A.0001.07.04.25.cdr

1A.2 Production Function

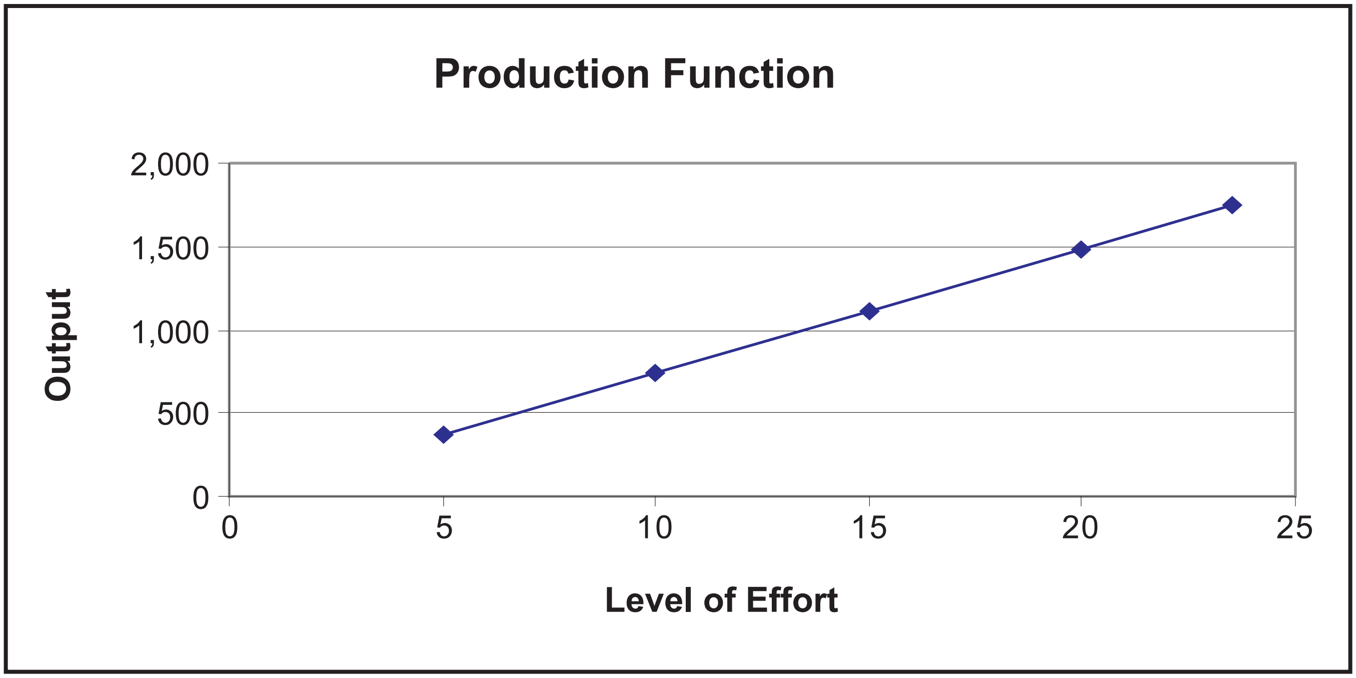
The relationship between the maximum output (production) that can be achieved from the application of a given input (level of effort)

The production function is the quantitative model of the production process, which can be any technology that will contribute to the desired output. A production function expresses the relationship between an organization's inputs and its outputs. It indicates, in mathematical or graphical form, what outputs can be obtained from various amounts and combinations of factor inputs. As described in the cost function above, the level of input is the size or amount of the technology that is employed. The level of output could be million gallons of overflows abated or tons of total suspended solids (TSS) removed. Figure 1A-2 is an example of a production function.

1A.3 Cost Benefit Relationship

The relationship between the output (benefit) of a technology and the cost to achieve that level of output

The cost benefit curve of a technology is the combination of the cost function and the production function. Using this relationship, the cost benefit of individual technologies can be compared. An example of a cost benefit relationship is shown in Figure 1A-3.



Preserving The Environment • Improving Water Quality

FIGURE 1A-2
EXAMPLE OF PRODUCTION FUNCTION
2020 STATE OF THE ART REPORT
4/25/07 SOAR_1A.0002.07.04.25.cdr

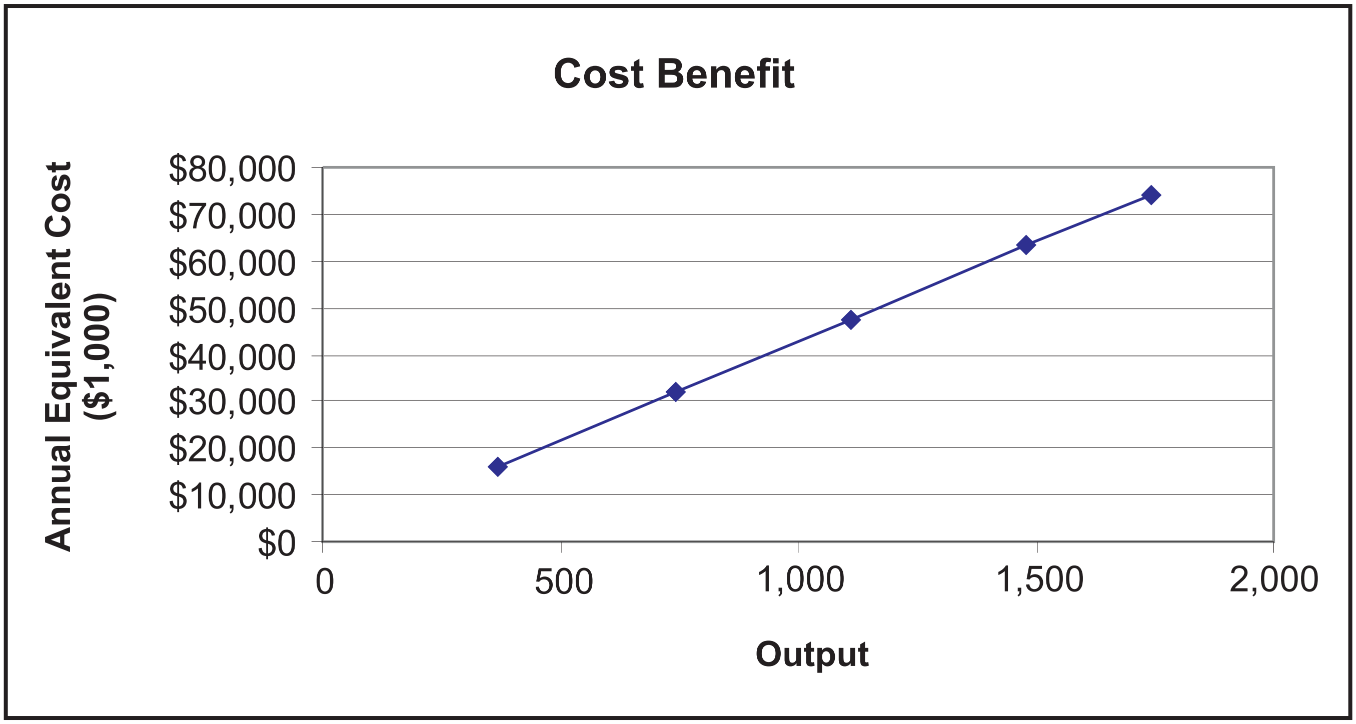


FIGURE 1A-3
EXAMPLE OF COST BENEFIT

1A.4 Interaction with other Control Technologies

The relationship between two or more technologies and the impacts on output or cost

Technologies can be independent of or dependent on other technologies, depending on factors such as the type, location, or the specific implementation of the technologies. When two technologies are employed in parallel or in series, there are three potential results:

1) They have no impact on each other
2) One technology increases the production of the other
3) One technology decreases the production of the other

For example, street sweeping in the combined sewer area does not directly impact the volume of tunnel storage required to reduce separate sewer overflows (SSOs). However, in a separate sewer area, street sweeping could reduce the discharge of TSS at a stormwater outfall equipped with ultraviolet (UV) disinfection, thereby allowing better light penetration and increased disinfection. Conversely, street sweeping an area tributary to a wet detention pond may decrease the production of the pond (i.e., pounds of material removed by the pond) because it would reduce the amount of material that would have settled out in the pond. Careful technical analysis is required when employing multiple technologies.

The remainder of this Appendix contains a more detailed explanation of the Production Function. This information is excerpted from the textbook, *The Foundations of Economic Analysis*(1).

The Production Function

"Until the laws of thermodynamics are repealed, I shall continue to relate outputs to inputs -- i.e. to believe in production functions."

(Paul A. Samuelson, *Collected Scientific Papers*, 1972: p.174)

Contents

(A) The Production Function

Let us begin with the *production function*, a function summarizing the process of conversion of factors into a particular commodity. We might propose a production function for a good y of the following general form, first proposed by Philip Wicksteed (1894):

$$y = f(x_1, x_2, ..., x_m)$$

which relates a single output y to a series of factors of production x_1, x_2, ..., x_m. Note that in writing production functions in this form, we are excluding *joint production*, i.e. that a particular process of production yields more than one output (e.g. the production of wheat grain often yields a co-product, straw; the production of omelettes yields the co-product broken egg shells). Using Ragnar Frisch's (1965) terms, we are concentrating on "single-ware" rather than "multi-ware" production.

For heuristic purposes, the production technology for the one-output/two-inputs case is (imperfectly) depicted in Figure 2.1. Output (Y) is measured on the vertical axis. The two inputs, which we call L and K which, for mnemonic purposes, can be called labor and capital,, are depicted on the horizontal axes. We ought to now warn that henceforth, throughout all our sections on the theory of production, *all capital is assumed to be endowed*, i.e. there are *no* produced means of production. The hill-shaped structure depicted in Figure 2.1 is the *production set*. Notice that it includes all the area *on* the surface and *in* the interior of the hill. The production set is essentially the set of technically feasible combinations of output Y and inputs, K and L.

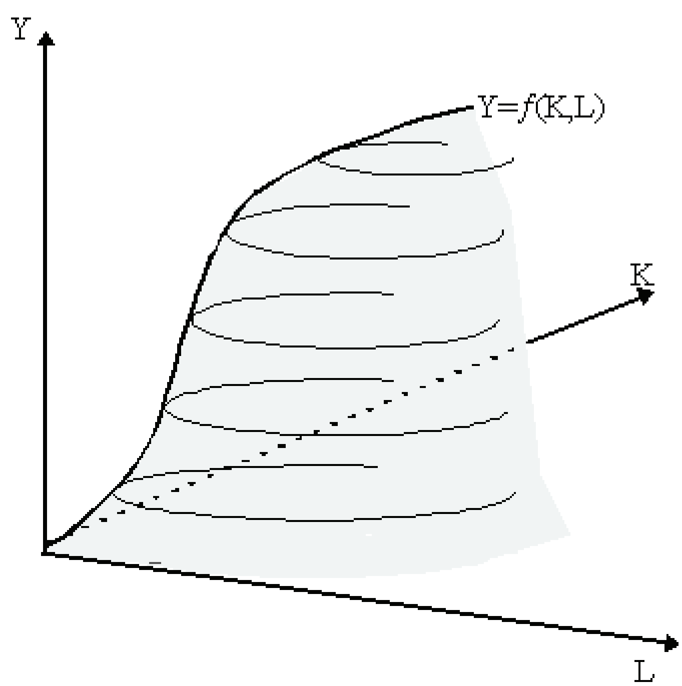


Figure 2.1 - Production function for one-output/two-inputs.

A *production decision* -- a feasible choice of inputs and output - is a particular point on or in this "hill". It will be "on" the hill if it is technically efficient and "in" the hill if it is technically inefficient. Properly speaking, the *production function* $Y = f(K, L)$ is *only* the surface (and not the interior) of the hill, and thus denotes the set of *technologically efficient* points of the production set (i.e. for a given configuration of inputs, K, L, output Y is the maximum feasible output).

Obviously, the hill-shape of the production function indicates that the more we use of the factors, the greater output is going to be (at least up to the some maximum, the "top" of the hill). The round contours along the production hill can be thought of as topographic contours as seen on maps and will serve as isoquants in our later analysis. The *slope* of the hill viewed from the origin captures the notion of returns to scale.

Throughout the next few sections, we shall be outlining the technical properties of the production function. The representation of production functions in the diagrammatic form of "hills" and the corresponding analysis of production theory in terms of isoquant contours, etc. was initiated by Vilfredo Pareto (1906) and much of the analysis of its technical properties was largely advanced by the "Paretian" school of Hotelling, Frisch, Samuelson, Hicks, Shephard, etc. between the 1930s and the 1950s. In the 1950s, the Neo-Walrasians (e.g. Koopmans, 1951; Debreu, 1959) approached the analysis of the technical properties of production in a somewhat different spirit. Specifically, instead of focusing on the "production function" and its derivatives as the Paretians had done, the Neo-Walrasians preferred to analyze it via vector space methods and convex analysis.

On a more formal note, we should outline the properties of the production function, as normally assumed by Neoclassical economists. Let there be m factors of production and let vector $x = (x_1, x_2, .., x_m)$ denote a bundle of factor inputs. We shall define an *input space* as the acceptable set of inputs for our economy. Commonly, a bundle of factor inputs x is deemed "acceptable" if every entry in that vector, i.e. the quantity of every factor, is a non-negative, finite real number. Thus, any input bundle x lies in R_+^m, the non-negative orthant of m-dimensional Euclidian space. Thus, R_+^m is our input space.

Let y be output, which is assumed to be a single, finite number, i.e. $y \in R$. Thus, a production function f maps acceptable input bundles to output values, i.e. $f: R_+^m \to R$. More specifically, $f(x)$ is the *maximum* output achievable for a given set of acceptable inputs, $x \in R_+^m$.

The following assumptions are often imposed on any generic production function $f : R_+^m \to R$

(A.1) $f(x)$ is finite, non-negative, real-valued and single-valued for all non-negative and finite x.

(A.2) $f(0, 0, .., 0) = 0$ (no inputs implies no output)

(A.3) If $x \geq x'$, then $f(x) \geq f(x')$ (monotonicity, i.e. an increase in inputs does not decrease output)

(A.4) f is *continuous* and twice-continuously *differentiable* everywhere in the interior of the production set.

(A.5) The set $V(y) = \{x \mid f(x) \geq y\}$ is a convex set (quasi-concavity of f)

(A.6) The set V(y) is closed and non-empty for any $y > 0$.

These assumptions will be clarified as we go on. For the moment, let us just make the following notes. Assumption (A.1) simply defines the production function as a well-defined function of inputs $f : R_+^m \rightarrow R$. Nothing new there. Assumption (A.2) simply establishes that one cannot produce something from nothing. This is somewhat self-evident, at least for economists. Obviously, in other walks of life, one can produce something without inputs (e.g. "nice thoughts" can just be, well, "thought up" without inputs), but most examples of these things are outside the realm of economics. The monotonicity assumption (A.3) is also straightforward: increasing inputs leads to an increase in output (or, more precisely, no decrease in output). Although common, we will have more to say on this later. Assumption (A.4) is made largely for mathematical ease; later on, we shall relax this assumption somewhat. Assumption (A.5), the quasi-concavity of the production function f , is the more interesting one. We shall have much more to say on this later. Finally, (A.6) is imposed as a mathematical necessity.

(B) Marginal Productivity

The assumptions given earlier imply that, for any given production function $y = f(x_1, x_2, .., x_m)$, it is a generally the case that, at least up to some maximum point:

$$\partial y/\partial x_i = f_i \geq 0$$

for all factor inputs i = 1, 2, ..., m. In other words, adding more units of any factor input will increase output (or at least not reduce it). This is the heart of assumption (A.3). However, it is also common in Neoclassical theory to also impose (A.5), i.e. to assume "quasi-concavity" of the production function. It is often the case in economics that the quasi-concavity assumption implies that:

$$\partial^2 y/\partial x_i^2 = f_{ii} < 0$$

for all i = 1, .., m, i.e. diminishing marginal productivity of ith factor.

It is worthwhile to spend a few moments on the diminishing marginal productivity assumption. This means more we add of a particular factor input, *all others factors remaining constant*, the less the employment of an additional unit of that factor input contributes to output as a whole. This concept performs the same function in production functions as diminishing marginal utility did in utility functions. Conceptually, however, they are quite distinct.

(i) The Law of Diminishing Returns

The idea of diminishing marginal productivity was simultaneously introduced for applications of factors to a fixed plot of land by T.R. Malthus (1815), Robert Torrens (1815), Edward West (1815) and David Ricardo (1815). It was applied more generally to other factors of production by proto-marginalists such as J.H. von Thünen (1826), Mountiford Longfield (1834) and Heinrich Mangoldt (1863). The apotheosis of the concept is found in the work of John Bates Clark (1889, 1891, 1899) and, more precisely, in Philip H. Wicksteed (1894). It was originally called the "*Law of Diminishing Returns*", although in order to keep this distinct from the idea of decreasing returns to scale, we shall refer to it henceforth as the "*Law of Diminshing Marginal Productivity*"

Let us first be clear about the definition of the *marginal productivity* of a factor. Letting Δx_i denote a unit increase in factor x_i, then the marginal product of that factor is $\Delta y/\Delta x_i$, i.e. the change in output arising from an increase in factor i by a unit. Mathematically, however, it is more convenient to assume that Δ x is infinitesimal. This permits us to express the marginal product of the factor x_i as the first partial derivative of the production function with respect to that factor -- thus the marginal product of the ith factor is simply $\partial y/\partial x_i = f_i$. If we do not wish to assume that factor units are infinitely divisible or if we do not assume that the production function is differentiable, we cannot express the marginal product mathematically as a derivative.

[We should note that both Carl Menger (1871) and John A. Hobson (1900, 1911) defined "marginal product" differently: rather than being the output gained by an enterprise from the addition of a factor unit, Hobson defined it as the output *lost* by the enterprise by the withdrawl of a factor unit. This caused a problem for the "adding up" issue in the marginal productivity theory of distribution, although, as was clarified later, when marginal product is not defined so discretely, it does not make a difference which measure we use. For a useful discussion of the dilemma involving the measurement of the "marginal unit", see the discussion in Fritz Machlup (1937). Finally, we must note that a far more novel and interesting definition of marginal productivity was introduced by Joseph M. Ostroy (1980, 1984) where the concept is redefined in terms of contributions to tradeable surpluses, and thus both widened and deepened in scope.]

However, assuming marginal products exist and are well defined, then why diminishing? Taking Clark's famous analogy:

> "Put one man only on a square mile of prairie, and he will get a rich return. Two laborers on the same ground will get less per man; and, if you enlarge the force to ten, the last man will perhaps get wages only."
>
> (J.B. Clark, 1890: p.304.)

The implication, then, is that as we increase the amount of labor applied to a particular fixed amount of land, each additional unit will increase total output but by *smaller and smaller* increments. When the field is empty, the first laborer has absolutely free range and produces as much as his body can reasonably do, say ten bushels of corn. When you add a second laborer to the same field, total output may increase, say to eighteen bushels of corn. Thus, the marginal product is eight.

Why? The basic idea is that by adding the second man, the field gets "crowded" and the men begin to get in each other's way. If that explanation does not seem credible, think of the units of labor in terms of labor-hours for a single man: in the first hour, a particular man produces ten; in the second hour he produces eight, etc. The diminution can be explained in this case as an "exhaustion" effect.

Taking another example, suppose we apply a man to a set of shoe-making tools and a given swathe of leather; let us say he can produce ten pairs of shoes in a day. Add a second man to this *without* adding more shoe-making tools or increasing the leather, and one can easily envisage that more shoes get made in a day, but that the work of the shoe-makers slows down as they pick up the same tools in an alternating sequence of turns and perhaps fight over them a bit.

For other factors, different stories are told. In Ricardo's original story, the land is subject to diminishing marginal returns because of the assumption that land has different degrees of fertility and the most fertile acres are used first, and the less fertile ones added later. We can conceive this more simply in that increasing the amount of land *without* increasing the amount of labor that works on it will lead to less output per worker.

However it is justified, many Neoclassical theorists basically accept diminishing marginal productivity *as an axiom* - "the diminishing marginal productivity of labor, when it is used in connection with a fixed amount of capital, is a universal phenomenon. This fact shows itself in any economy, primitive or social." (Clark, 1899: p.49). However much early economists tried to claim it to be a natural law, this "axiom" turns out to be closer to a rather debatable assumption (cf. Karl Menger, 1954).

Nonetheless, it is important to clearly note a few matters in relation to this. Firstly, the idea that marginal product is always diminishing can be disputed (and will be disputed). Francis A. Walker (1891) took J.B. Clark to task for not recognizing the possibility of increasing marginal productivity (albeit, see Clark (1899: p.164)).

Secondly, as Pareto (1896, 1902) was quick to point out, it is not always true that if one adds a unit of a factor to an existing production process, output will increase. "If a pit has to be dug, the addition of one more man will make little difference to the day's output unless you give the man a spade" (Cassel, 1918: p.179). This difficulty is even more clear if we see the problem in terms of the marginal product of capital: if a pit has to be dug, the addition of one more spade will make no difference to output unless you add a man to use it. Thus, one must be very careful when pronouncing the idea of marginal productivity since we may need to produce in fixed, constant factor proportions.

Thirdly, it is important to underline that the marginal product is not, properly speaking, the contribution of the marginal unit *by itself*. Some commentators (e.g. E.v. Böhm-Bawerk, 19??; cf. Robertson, 1931) seem to have gone on to make arguments that seem to imply, in the context of our example, that the second man produces *eight* bushels of corn. Of course, this is not necessarily true. The second man may very well produce nine or ten or eleven and *still* the total output increases only to eighteen because the *first* man reduces his output to nine, or eight or seven in the presence of the second. In our example, output increases from ten bushels to eighteen bushels when one adds the second man *not* because the second man only adds eight, but rather because his presence on the field makes the situation such that the total output of *both* men is eighteen. Notice that the contribution *per man* is reduced: the average product is actually nine. This may very well be how much each of the two laborers contributes. But this is not what interests us: what we wish to note is that by *adding* the second man, output was increased by *eight*. *Thus*, the marginal product of the second man is eight. But his *actual* contribution may be very different than this.

Finally, and above everything, it is very important to note that in deriving the marginal product of a factor, we are holding *all other factors fixed*. Specifically, in our earlier example, labor varied and land (and indeed all other factors) was fixed. Thus, diminishing marginal productivity has *nothing* to do with "returns to scale", i.e. the increase in output when we increase *all* factors. If we increased *both* land and labor in our example, then there might very well be no reduction in output per man (indeed, there is actually no reason for it, but we shall return to this later).

(ii) The Law of Variable Proportions

Marginal productivity is *not* obvious in the production function $Y = f(L, K)$ in Figure 2.1 as both inputs are varying there. We must first fix one of the factors and let the other factor vary. This is shown in Figure 2.2, by the "reduced" production function $Y = f(L, K_0)$, where only labor (L) varies while capital is held fixed at K_0. To obtain this from the former, we must figuratively "slice" the hill in Figure 2.1 vertically at the level K_0. Thus, Figure 2.2, which represents the reduced production function $Y = f(L, K_0)$, is a vertical section of the hill in Figure 2.1. A reduced production function where all factors but one are held constant are often referred to as the "*total product*" curve.

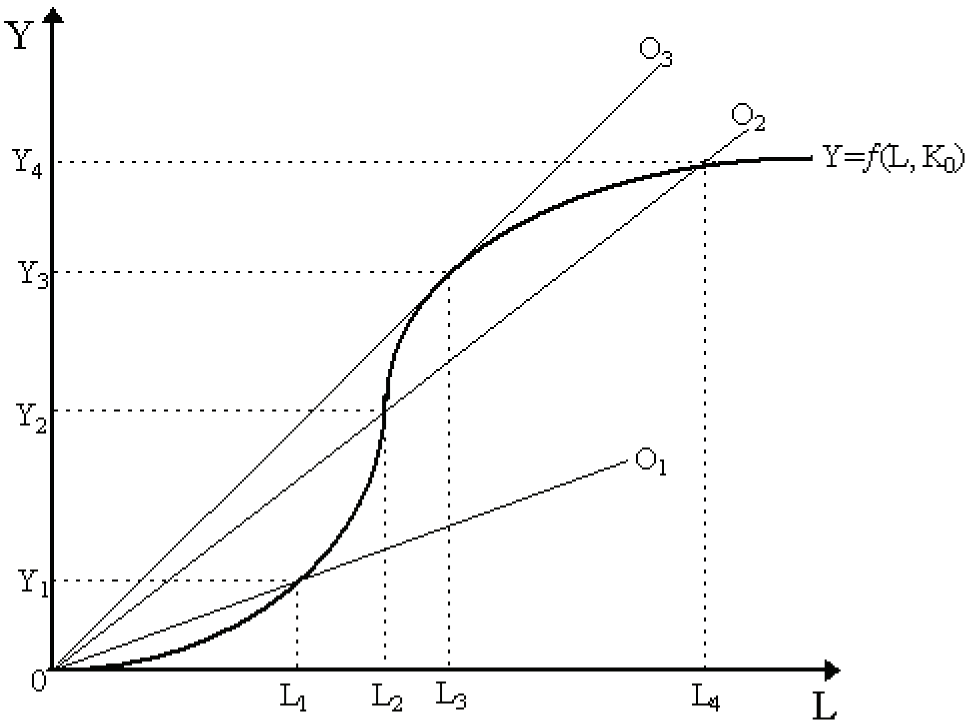


Figure 2.2 - Total Product Curve

The total product curve in Figure 2.2 can be read in conjunction with the average and marginal product curves in Figure 2.3. The total product curve is originally due to Frank H. Knight (1921: p.100), and much of the subsequent analysis is due to him and John M. Cassels (1936). Although both these sets of curves have long been implicit in much earlier discussions (e.g. Edgeworth, 1911), average and marginal products were confused by early Neoclassicals with surprising frequency. The particular shape of the total product curve shown in Figure 2.2 exhibits what has been baptized by John M. Cassels (1936) as *the Law of Variable Proportions* -- effectively what Ragnar Frisch (1965: p.120) quirkily renamed the *ultra-passum law* of production.

The *marginal product* of the factor L is given by the *slope* of the total product curve, thus $MP_L = \partial Y/\partial L = df(L, K_0)/dL$. As we see, at low levels of L up to L_2 in Figure 2.2, we have *rising* marginal productivity of the factor. At levels of L above L_2 we have *diminishing* marginal productivity of that factor. Thus, marginal productivity of L reaches its maximum at L_2. We can thus trace out a marginal product of L curve, MP_L, in Figure 2.3. The labels there correspond to those of Figure 2.2. Thus the MP_L curve in Figure 2.3 rises until the inflection point L_2, and falls after it. It becomes negative after L_5 - which would be equivalent to the "top" of the reduced production function, what Frisch (1965: p.89) calls a "strangulation point". A negative marginal product is akin to a situation when one adds the fiftieth worker to a field whose only accomplishment is to get in everyone else's way - and thus does not increase output at all but actually reduces it.

The slope of the different *rays* through the origin (O_1, O_2, O_3, etc.) in Figure 2.2 reflect *average products* of the factor L, i.e. $AP_L = Y/L$. The steeper the ray, the higher the average product. Thus, at low levels of output such as Y_1, the average product represented by the slope of O_1 is rather low, while at some levels of output such as Y_3, the average product (here the slope of O_3) is much higher. Indeed, as we can see, average product is at its highest at Y_3, what is sometimes called the *extensive margin* of production. Notice that at Y_2 and Y_4 we have the same average product (i.e. the ray O_2 passes through both points). The average product curve AP_L corresponding to Figure 2.2 is also drawn in Figure 2.3.

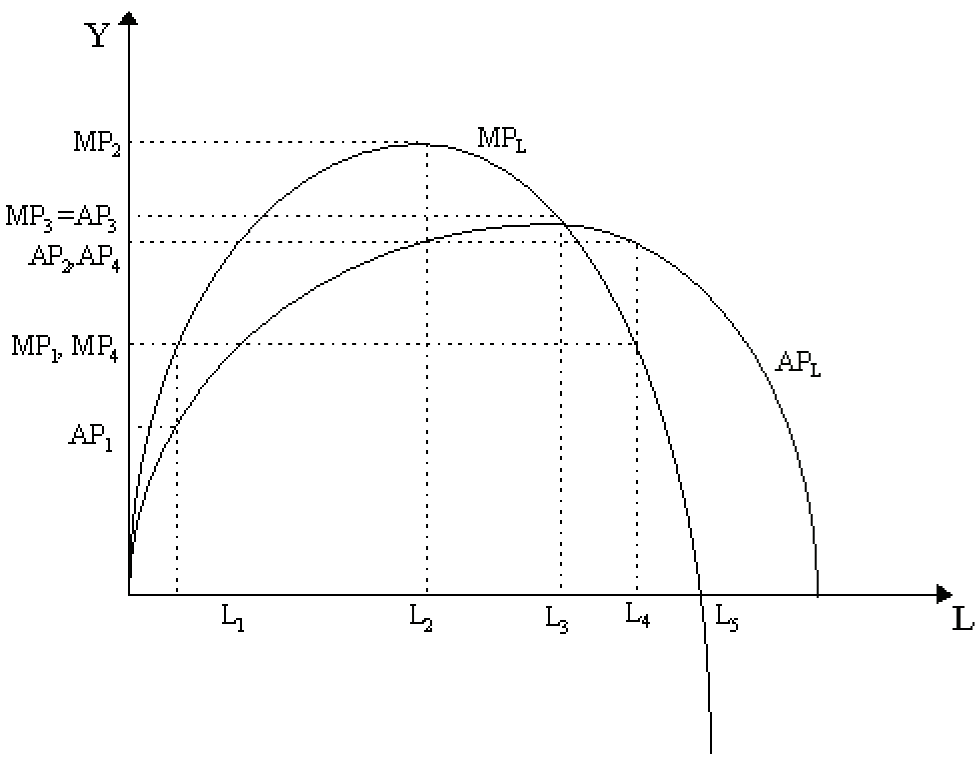


Figure 2.3 - Marginal Product and Average Product curves

As we can see in Figure 2.2, the slope of the total product curve is equal to the slope of the ray from the origin at L_3, thus average product and marginal product are equal at this point (as shown in Figure 2.3). We also know that as the ray from the origin associated with L_3 is the highest, thus average product curve intersects the marginal product curve, $MP_L = AP_L$, exactly where the average product curve is at its maximum. Notice that at values below L_3, $MP_L > AP_L$, marginal product is greater than average product whereas above L_3, we have the reverse, $MP_L < AP_L$. We shall make use of these results later on.

As we can see in Figure 2.3, it seems that we can have increasing as well as diminishing marginal productivity of labor, as suggested by Walker (1891) and finally acknowledged by Clark (1899: p.164). However, we have already gone a long way in arguing for diminishing marginal productivity that it seems that we must be excluding points where there is rising marginal productivity, i.e. those points to the left of L_2.

How might such a restriction be justified? In effect, the argument is that in situations of increasing marginal productivity, one can always *discard* factors and *increase* output (cf. F.H. Knight, 1921: p.100-104; G. Cassel, 1918: p.279; J.M. Cassels, 1936; A.P. Lerner, 1944: p.153-5). Consider the following example. Assume we have an acre of ripened land to which we are going to apply various quantities of workers. The one lonely worker produces 10 bushels of wheat; two workers will produce 22 bushels of wheat; three workers will produce 36 bushels.

Thus, we see:

Qty. of Labor	Total Product	Average Product	Marginal Product
One Laborers	10	10	10
Two Laborers	22	11	12
Three Laborers	36	12	14

Thus, there is increasing average product and increasing marginal product of labor in this example. Why? One can think of it as follows. When we apply one lonely worker to an entire acre of ripened land, his running around the entire acre trying to harvest it will produce a lower average product than if we had three workers, each working a third of the field by himself.

This should already reveal why we would never see a situation of increasing average product. Basically, when we are faced with a situation of a single worker on an acre of land, why should we force him to work on the entire acre and only produce 10 units of output? Average product (and total product) would be *higher* if instead of forcing that single worker to try to harvest the entire acre, we let him confine himself to a third of that acre, and let the other two-thirds of the plot lie untouched. In this case, the average product of the single worker is as it would have been had there been three harvesters, i.e. 12 units of output. In other words, in situations of rising average and marginal product, total output is *increased* by discarding two-thirds of the land! Thus, situations of increasing marginal productivity will simply never be seen.

Of course, this logic is not unassailable. While the idea may apply naturally to some cases, it can be questioned in cases where division of labor is crucial as, say, we might have in an automobile factory. Suppose that the average productivity of a worker is highest when there are twenty men working on a factory floor, each worker specializing in fitting a special part of the automobile. We cannot subsequently do the same operation we did before. In other words, we cannot remove nineteen men and let 19/20ths of the car remain unbuilt. The only remaining man, whose productivity was highest when he only fitted wheels on axles, will not yield *any* output if he is permitted to perform only his specialized task bereft of the other nineteen men. Instead of having cars as output in that case, we would have axles-with-wheels.

Consequently, we see that in order to produce any cars whatsoever, the lone man must be forced to perform *all* the tasks, not only the fitting of wheels on axles. If this is true, his productivity by himself, where product is measured in number of cars produced rather than axles-with-wheels, will be *lower* than if he worked together with his nineteen colleagues.

The automobile case shows an example of *indivisibility* in production, a traditional explanation of increasing marginal productivity (e.g. Edgeworth, 1911; Lerner, 1944). Production is *divisible* if it "permits any particular method of production, involving certain proportions between factors and products, to be repeated in exactly the same way on larger or on a smaller scale." (Lerner, 1944: p.143). In other words, in a perfectly divisible world, there cannot be *changes* in method when increasing or decreasing the scale of production. In our automobile example we have indivisibility: when we remove the nineteen men, the remaining man who previously only placed axles on wheels must *change* his method and do all the tasks in the construction of the automobile. In contrast, our agricultural example was divisible: a laborer working exclusively on his portion of the field will not change his method of harvesting that third of the field when the other laborers on the other the remaining two-thirds of the field are removed.

In sum, increasing marginal productivity, especially in cases where specialization is vital, can be ostensibly encountered in the real world where there are indivisibilities in production. Nevertheless, much of the Neoclassical work on the production function omits this. This is, as noted earlier, is often taken axiomatically, but the question of whether one finds it an acceptable assumption is largely an empirical one.

(C) Isoquant Analysis

The contours along the production "hill" in Figure 2.1 are the isoquants shown in Figure 2.4. A particular isoquant denotes the combinations of factors K and L which produce the same quantity of output. As we are assuming factors K and L are continuously substitutable (on which we will have more to say later), then every point on a particular isoquant represents a particular feasible *technique*, or factor combination, that can be used to produce a *particular* level of output. The isoquants play the same topographic role to the production "hill" as indifference curves played in the the "utility hill". As the isoquants ascend to the northeast, the amount of output produced increases, thus $Y' < Y^* < Y''$.

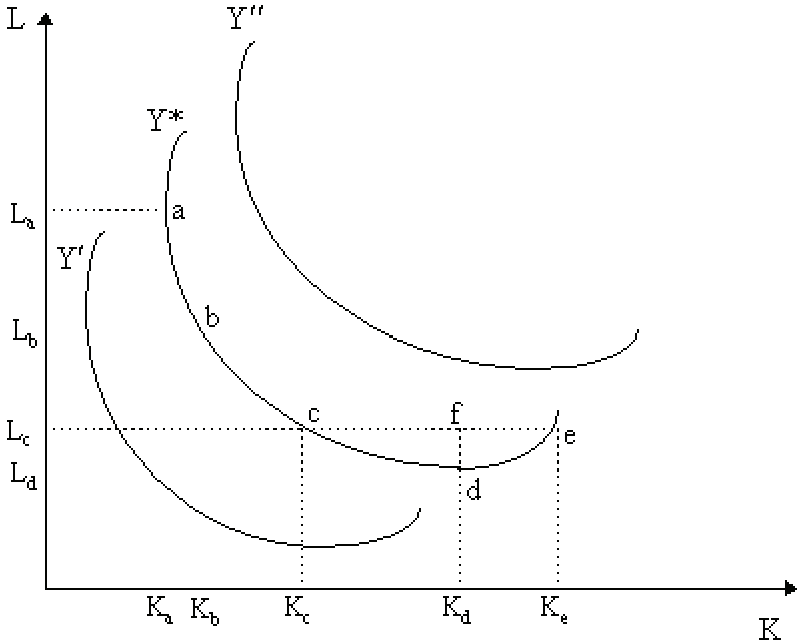


Figure 2.4 - Isoquants

It is an elementary matter to derive the slope of an isoquant. For our two-factor case, we had a production function $Y = f(L, K)$. Now, for the production of a given fixed quantity of output (call it Y^*), it follows that $Y^* = f(L, K)$. This is the formula for a particular isoquant. Totally differentiating this:

$$dY^* = f_L dL + f_K dK$$

where $f_L = \partial Y/\partial L$ and $f_K = \partial Y/\partial K$ are the marginal products of labor and capital respectively, evaluated around Y^*. Since on any isoquant, output is fixed at Y^*, then $dY^* = 0$. This implies that $f_L dL = -f_K dK$, or simply:

$$-dL/dK|_{Y^*} = f_K/f_L$$

The term on the left is the negative of the slope of the isoquant corresponding to output level Y^*. This is known as the *marginal rate of technical substitution* (MRTS), i.e. the rate at which capital can be susbstituted for labor while holding output constant along an isoquant. (note that dL/dK by itself is already negative, thus the MRTS will be a positive number). Provided our isoquants are smoothly differentiable, we will be able to define the MRTS at any point in Figure 2.4. Thus, the MRTS depends not only on the level of output (which isoquant we are on), but also the amounts of capital and labor (where on the isoquant we are).

The equality of the MRTS with the ratio of marginal products of capital and labor, f_K/f_L, is a fundamental feature of production theory and helps us capture the concept of diminishing marginal productivity to a factor. In Figure 2.4, on isoquant Y*, as we move from point a to b to c to d, we are moving towards greater employment of K and less employment of L to produce a given level of output Y*, thus we are moving from *labor-intensive* techniques (i.e. low capital-labor ratios) towards *capital-intensive* techniques (high capital-labor ratios). Notice also that the isoquant becomes flatter as we move from a to d, thus the marginal rate of technical substitution is *higher* at a than at d, i.e. $MRTS_a > MRTS_b > MRTS_c > MRTS_d$. Thus, there is *diminishing* marginal rates of technical substitution as we move from a towards d.

Notice that this declining MRTS arises because of the *convexity* of the isoquants. As we can notice, the declining MRTS effectively can capture something akin to (but not exactly) of the assumption of diminishing marginal productivity to a factor we spoke of earlier. Compare only the points b and c. As $MRTS_b > MRTS_c$, then $f_K/f_L|_b > f_K/f_L|_c$. But point b represents a lower capital-labor ratio than point c, i.e. $K/L|_b < K/L|_c$. Thus, we can interpret the declining MRTS as saying that as we move from lower capital-intensity to higher capital intensity (b to c), the marginal product of capital *decreases*. Reciprocally, as we move from higher labor-intensity to lower labor-intensity (b to c), the marginal product of labor *increases*.

Note that we cannot *derive* diminishing MRTS from the assumption of diminishing marginal productivity of factors. To see this, consider the production function, $Y = f(K, L)$. The MRTS at any point is f_K/f_L. In order to have *diminishing* MRTS, then it must be that dMRTS/dK < 0. But as $dMRTS/dK = d(f_K/f_L)/dK$ and f_K and f_L vary with K and L, then we must take total derivatives. Thus:

$$dMRTS/dK = d(f_K/f_L)/dK$$

$$= [(f_{KK} + f_{KL}\cdot dL/dK)f_L - (f_{LK} + f_{LL}\cdot dL/dK)f_K]/f_L^2$$

as $dL/dK = -f_K/f_L$ along any isoquant and given that $f_{KL} = f_{LK}$ by Young's Theorem, then:

$$dMRTS/dK = [f_{KK}\, f_L - 2f_{LK}\, f_K + f_{LL}\cdot f_K^2/f_L]/f_L^2$$

or simply:

$$dMRTS/dK = [f_{KK} f_L^2 - 2 f_{LK} f_K f_L + f_{LL} \cdot f_K^2]/f_L^3$$

Now, by assumption, f_K, $f_L > 0$ and, by diminshing marginal productivity, $f_{KK} < 0$ and $f_{LL} < 0$. This is obviously not sufficient to determine the sign of dMRTS/dK. Specifically, the numerator will only be negative if, in addition, we assume that $f_{LK} > 0$, and the theory of diminishing marginal productivity implies no such thing. Thus, a diminishing MRTS is, in itself, an separate assumption.

We should note, however, that *not* every point along the isoquant is relevant. The isoquants, after all, are contours of our "production" hill and thus are actually "circular". This is captured in Figure 2.5, where we show the isoquants in their full topographic glory as a horizontal section of the production hill of our earlier Figure 2.1. Notice that the isoquant labels represent increasing output levels, $Y < Y' < Y'' < Y'''$, etc. The "top of the hill", the highest output achievable, is represented by point M in the center, achieved by factor combination L_M and K_M. Notice that if we are the top of the hill, *if* we increase factor inputs (above K_M or L_M), output will actually *decline*.

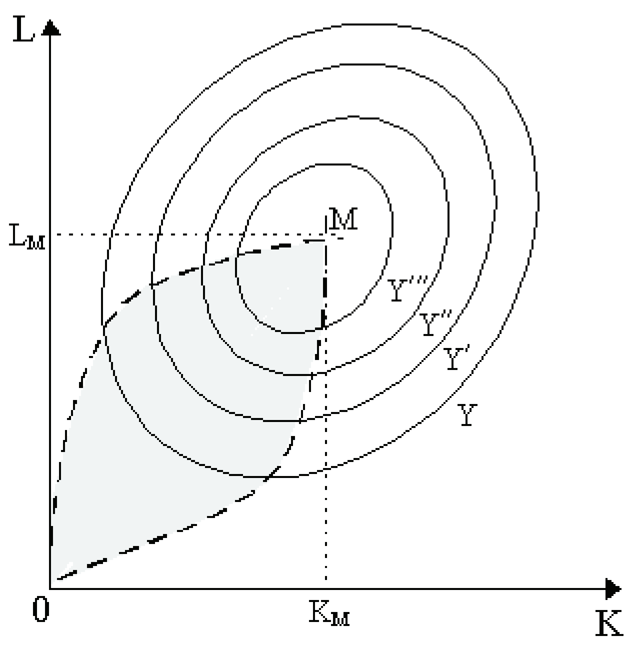


Figure 2.5 - Isoquants with Ridge Lines

We have also added dashed "ridge lines" to the topographic map in Figure 2.5. Only those points *within* the ridge lines, in the lightly shaded region, are of economic relevance. To see why, return to Figure 2.4 and notice that at point a, the isoquant has a vertical slope and a point d, the isoquant has a horizontal slope. Thus, $MRTS_a = f_K/f_L|_a = \infty$ and $MRTS_d = f_K/f_L|_d = 0$. But our isoquants seem to continue beyond them, yet we assert that points beyond them are economically irrelevant.

Why? Consider a factor combination such as at point e in Figure 2.4. Obviously, here, the slope of the isoquant is positive, i.e. $dL/dK|_e > 0$, which implies, in turn, that $MRTS_e = f_K/f_L < 0$, thus the marginal product of one of the factors is negative. This violates the first assumption we made about the production function: namely, that $f_i > 0$ for all i, i.e. increasing the employment of any factor in a production process will always increase output. Thus, we ought to exclude all regions where marginal products are negative.

Is this assumption reasonable? Well, notice that at point e, we are employing factors K_e and L_e to produce output level Y*. Yet, we could *decrease* the amount of capital employed to K_d and leave labor at L_e in order to achieve a combination at point f. But notice that as point f is *above* the isoquant Y*, it effectively represents a *higher* level of output. Thus, if we are at a point such as e, then by *reducing* factor inputs we can *increase* output: such factor combinations are therefore not "economical". Consequently we can rule out point e - and, indeed, all factor combinations on the isoquant Y* beyond d in Figure 2.4. Similarly, we exclude points on the isoquant beyond point a for the same reason.

The "ridge lines" drawn in Figure 2.5, pass through limiting points of the various isoquants akin to points a and d in Figure 2.4. In other words, at any point on the upper ridge line, MRTS = ∞ for the relevant isoquant, while at any point on the lower ridge line, MRTS = 0 for the relevant isoquant. Thus, we *exclude* all regions above the upper ridge line and below the lower ridge line as economically irrelevant. Only the lightly shaded area in Figure 2.5 is "relevent". Notice that the ridge lines meet at point M, the "top" of the production "hill".

On a more formal note, we should connect the quasi-concavity of the production function to the convexity of the isoquants in general. A function f is *quasi-concave* if, for any two input bundles $x, x' \in R_+^m$

if $f(x) \geq f(x')$, then $f(\lambda x + (1-\lambda)x') \geq f(x')$ for any $\lambda \in (0, 1)$.

In other words, the output produced from a convex combination (i.e. weighted average) of two bundles of inputs is at least as great as the smaller of the two outputs produced using only one or the other input bundles. A special case of quasi-concave function is simply a *concave* function, namely for any two input bundles x, x′ ∈ R_+^m:

$$f(\lambda x + (1-\lambda)x') \geq \lambda f(x) + (1-\lambda)f(x') \text{ for any } \lambda \in (0, 1).$$

which states that the output produced from a convex combination of inputs is at least as great as the convex combination of the outputs produced by the input bundles independently.

The definition of quasi-concavity we used in (A.5) states that V(y) = $\{x \mid f(x) \geq y\}$ is convex. In other words, a function is quasi-concave if the *upper contour set* V(y) is convex. As we see, this "upper contour set" V(y) is merely the isoquant defined by y and the area above that isoquant. This is illustrated in Figure 2.6. Y* is the isoquant of relevance, thus V(Y*), the shaded area, is the upper contour set.

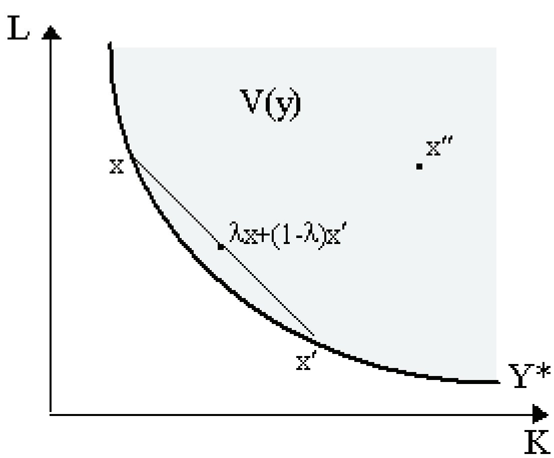


Fig. 2.6 - Upper Contour Set and Convexity

[It is worth pointing out now what the assumption (A.6) on production meant. Effectively, if V(y) has an interior point, then (A.6) intuitively states that for any upper contour set V(y), there is a factor bundle "inside" it (e.g. x′ ′ in Figure 2.6). In order to guarantee (A.6), it is common practice in Neo-Walrasian production theory to impose the assumption of *free disposibility* in production. What this assumption states, effectively, is that if one can produce a particular level of output y with input bundle x, then one can produce an amount of output which is *less* than y with that same input bundle, x or, equivalently, one can produce the *same* level of output y with *more* inputs. The idea is that the producer just costlessly throws away the "extra" that has been produced or, equivalently just throws away some of the extra factors rather than using them. Thus, in Figure 2.6, x'' can be used to produce y if the extra amount normally produced by x'' is "thrown away" or if the extra factors are just left unused by the producer. The free disposal assumption, then, is meant to guarantee at least the technical possibility of "inefficient" production, and thus interior points to the production set and the input requirement sets. However extensively used in production theory, it might be regarded nonetheless as somewhat stronger assumption than one might wish for and, thus many economists have endeavoured to dispose of it.]

It can be easily proven that if a function is quasi-concave in the sense defined earlier, then its upper contour set is necessarily convex. To see this intuitively, let x and x' be two points on the same isoquant, thus $f(x) = f(x')$, as shown in Figure 2.6. Thus, quasi-concavity implies that the output produced by any convex combination of the two points x and x' is greater than the output produced by either point individually. In other words, the convex combination of two points on the same isoquant will lie on a *higher* isoquant. In Figure 6, we see that $\lambda x + (1-\lambda)x'$ is indeed *within* V(y), thus it will lie on a higher isoquant. This implies precisely the convex shape of the isoquant curves which implies, in turn, the convexity of the upper contour set V(y).

A final characterization of quasi-concavity makes use of the bordered Hessian matrix. The bordered Hessian matrix of a twice-continuously differentiable function $y = f(x_1, x_2, .., x_m)$ is defined as follows:

$$B = \begin{matrix} 0 & f_1 & f_2 & \cdots & f_m \\ f_1 & f_{11} & f_{12} & \cdots & f_{1m} \\ f_2 & f_{21} & f_{22} & \cdots & f_{2m} \\ : & \ldots & \ldots & \ldots & \ldots \\ f_m & f_{m1} & f_{m2} & \cdots & f_{mm} \end{matrix}$$

where f_i is the first partial derivative of the production function with respect to factor x_i and f_{ij} are the second derivatives, all evaluated at a particular factor combination x.

If the production function is quasi-concave, then we know that the bordered Hessian of that function evaluated at any input bundle $x \in R_+^m$ will be negative semi-definite, i.e. its principal leading minors will alternate in sign. Specifically:

$$\begin{vmatrix} 0 & f_1 \\ f_1 & f_{11} \end{vmatrix} \leq 0$$

$$\begin{vmatrix} 0 & f_1 & f_2 \\ f_1 & f_{11} & f_{12} \\ f_2 & f_{21} & f_{22} \end{vmatrix} \geq 0$$

... etc.

Notice that the very first principal minor implies that $-f_1^2 \leq 0$, which is true whether $f_1 \geq 0$ or $f_1 \leq 0$, so quasi-concavity does not rule out negative marginal products. The second principal minor implies:

$$-f_1[f_1 f_{22} - f_2 f_{21}] + f_2[f_1 f_{12} - f_2 f_{11}] \geq 0$$

as, byYoung's Theorem, $f_{21} = f_{12}$, this is reducible to:

$$2f_1 f_2 f_{21} - f_1^2 f_{22} - f_2^2 f_{11} \geq 0.$$

Now, even if $f_1 \geq 0$ and $f_2 \geq 0$ by assumption, there is little implied by this condition. In other words, quasi-concavity of the production function is *not* sufficient to guarantee *diminishing* marginal productivities, i.e. $f_{11} \leq 0$, $f_{22} \leq 0$, etc.

In contrast, *concavity* of the production function is, indeed, enough to yield diminishing marginal productivity. We can verify this by just examining the Hessian for the production function. This is:

$$f_{11} \quad f_{12} \quad \cdots \quad f_{1m}$$

$$H = \begin{matrix} f_{21} & f_{22} & \cdots & f_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ f_{m1} & f_{m2} & \cdots & f_{mm} \end{matrix}$$

where, note, the border is omitted. Concavity implies that the Hessian matrix must be negative semi-definite, by which we mean that the principle leading minors alternate in sign. This implies that:

$f_{11} \leq 0$

$$\begin{matrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{matrix} \geq 0$$

... etc.

Notice that the very first principal leading minor states that $f_{11} \leq 0$, i.e. negative marginal productivity for input 1. As we can order inputs anyway we wish, then this effectively generalizes to stating that every factor exhibits diminishing marginal productivity, i.e. $f_{ii} < 0$ for all i = 1, 2, .., m. Thus, while quasi-concavity cannot guarantee diminishing marginal productivity to factor inputs, concavity does indeed guarantee it. However, we should note that there are special cases when quasi-concavity of the production function guarantees diminishing marginal productivity - namely, under constant returns to scale. We will have more to say on this later.

Theories of Production – Selected References

A. Aftalion (1911) "Les Trois Notions de la Productivité et les Revenues", *Revue d'Economie Politique*. Vol. 25, p.145-84.

R.G.D. Allen (1938) *Mathematical Analysis for Economists*. New York: St. Martin's.

K.J. Arrow, H.B. Chenery, B.S. Minhas and R.M. Solow (1961) "Capital-Labor Substitution and Economic Efficiency", *Review of Economics and Statistics*, Vol. 43, p.225-50.

E. Barone (1895) "Sopra un Libro del Wicksell", *Giornale degli Economisti*, Vol. 11, p.524-39.

E. Barone (1896) "Studie sulla Distribuzione", *Giornale degli Economisti*, Vol. 12, p.107-55; 235-52.

W.J. Baumol (1958) "Activity Analysis in One Lesson", *American Economic Review*, Vol. 68, p.665-72.

W.J. Baumol (1961) *Economic Theory and Operations Research*. 1976 edition, Englewood Cliffs, NJ: Pretntice-Hall.

C. Blackorby and R.R. Russell (1981) "The Morishima Elasticity of Substitution: Symmetry, consistency, separability and its relationship to the Hicks and Allen elasticities", *Review of Economic Studies*, Vol. 48, p.147-58.

C. Blackorby and R.R. Russell (1989) "Will the Real Elasticity of Substitution Please Stand Up? A comparison of the Allen/Uzawa and Morishima elasticities", *American Economic Review*, Vol. 49 (4), p.882-88.

M. Bronfenbrenner (1960) "A Note on Relative Shares and the Elasticity of Substitution", *Journal of Political Economy*, Vol. 68, p.284-7.

D.J. Brown (1991) "", in W. Hildenband and H. Sonnenschein, editors, *Handbook of Mathematical Economics, Vol. IV*. Amsterdam: North-Holland.

S. Carlson (1939) *A Study on the Pure Theory of Production*. 1956 edition, New York: Augustus M. Kelley.

G. Cassel (1918) *The Theory of Social Economy*. 1932 edition, New York: Harcourt, Brace..

J.M. Cassels (1936) "On the Law of Variable Proportions", in *Explorations in Economics in honor of Frank Taussig*. New York: McGraw-Hill. Reprinted in American Economic Association, *Readings in the Theory of Income Distribution*. New York: Blakiston.

E.H. Chamberlin (1948) "Proportionality, Divisibility and Economies of Scale", *Quarterly Journal of Economics*

R.G. Chambers (1988) *Applied Production Analysis: A dual approach*. Cambridge, UK: Cambridge University Press.

D.G. Champernowne (1953) "The Production Function and the Theory of Capital: A comment", *Review of Economic Studies*, Vol. 21 (1), p.112-35.

S.J. Chapman (1906) "The Remuneration of Employers", *Economic Journal*, Vol. 16, p.523-28.

J.B. Clark (1888) "Capital and Its Earnings", *Publications of the American Economic Association*, 3 (2).

J.B. Clark (1889) "Possibility of a Scientific Law of Wages", *Publications of the American Economic Association*, Vol. 4 (1)

J.B. Clark (1891) "Distribution as Determined by a Law of Rent", *Quarterly Journal of Economics*, Vol. 5, p.289-318.

J.B. Clark (1893) "The Genesis of Capital", *Yale Review*, Vol. 2, p.302-15.

J.B. Clark (1895) "The Origin of Interest", *Quarterly Journal of Economics*, Vol. 9, p.257-78.

J.B. Clark (1899) *The Distribution of Wealth: A theory of wages, interest and profits*. 1927 edition, New York: Macmillan.

J.B. Clark (1901) "Wages and Interest as Determined by Marginal Productivity", *Journal of Political Economy*, Vol. 10, p.

C.W. Cobb and P.H. Douglas (1928) "A Theory of Production", *American Economic Review*, Vol. 18, p.139-65.

B. Cornet (1982)

H.J. Davenport (1913) *The Economics of Enterprise*.

Gerard Debreu (1959) *Theory of Value: An axiomatic analysis of economic equilibrium*. New York: Wiley.

W.E. Diewert (1974) "Applications of Duality Theory", in M.D. Intriligator and D.A. Kendrick, editors, *Frontiers of Quantitative Economics, Vol. II*. Amsterdam: North-Holland.

W.E. Diewert (1982) "Duality Approaches to Microeconomic Theory", in K.J. Arrow and M.D. Intriligator, editors, *Handbook of Mathematical Economics, Vol. II*. Amsterdam: North-Holland.

R. Dorfman (1953) "Mathematical or `Linear' Programming: A nonmathematical exposition", *American Economic Review*, Vol. 43 (5), p.797-825.

R. Dorfman, P.A. Samuelson and R.M. Solow (1958) *Linear Programming and Economic Analysis*. 1987 reprint, New York: Dover.

P.H. Douglas (1934) *The Theory of Wages*. New York: Macmillan.

P.H. Douglas (1948) "Are There Laws of Production?", *American Economic Review*, Vol. 38, p.1-41.

P.H. Douglas (1967) "The Cobb-Douglas Production Function Once Again: Its history, its testing and some empirical values", *Journal of Political Economy*, Vol. 84 (5), p.903-16.

J.H. Drèze (1985) "Uncertainty and the Firm in General Equilibrium Theory", *Economic Journal*, 1985,

F.Y. Edgeworth (1904) "The Theory of Distribution" *Quarterly Journal of Economics*, Vol. 18, p. 149-219.

F.Y. Edgeworth (19??) "The Laws of Increasing and Diminishing Returns", *??.*

F.Y. Edgeworth (1911) "Contributions to the Theory of Railway Rates, I & II", *Economic Journal*, Vol. 21, p.346-71, 551-71.

C.E. Ferguson (1969) *The Neoclassical Theory of Production and Distribution*. Cambridge, UK: Cambridge University Press.

R. Frisch (1950) "Alfred Marshall's Theory of Value", Vol. 65 (4), p.495-524.

R. Frisch (1965) *Theory of Production*. Dordrecht: D. Reidel.

M. Fuss and D. McFadden (1978), editors, *Production Economics: A dual approach to theory and application*. Amsterdam: North-Holland.

A.W. Flux (1894) "Review of Wicksteed's Essay", *Economic Journal*, Vol. 4, p.305-8.

D. Gale (1960) *The Theory of Linear Economic Models*. New York: McGraw-Hill.

G.C. Harcourt (1972) *Some Cambridge Controversies in the Theory of Capital*. Cambridge, UK: Cambridge University Press.

J. Hicks (1931) "The Theory of Uncertainty and Profit", *Economica*, Vol. 11, p.170-89.

J. Hicks (1932) *The Theory of Wages*. 1963 edition, London: Macmillan.

J. Hicks (1934) "Marginal Productivity and the Principle of Variation", *Economica*,

J. Hicks (1936) "Distribution and Economic Progress: A revised version", *Review of Economic Studies*, Vol. 4 (1), p.1-12.

J. Hicks (1939) *Value and Capital: An inquiry into some fundamental principles of economic theory*. 1946 Edition, Oxford: Clarendon Press.

J. Hicks (1965) *Capital and Growth*. Oxford: Clarendon.

J.A. Hobson (1891) "The Law of the Three Rents", *Quarterly Journal of Economics*, Vol. 5, p.263-88.

J.A. Hobson (1900) *The Economics of Distribution*. New York: Macmillan.

J.A. Hobson (1909) *The Industrial System*. London: Longmans.

H. Hotelling (1932) "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions", *Journal of Political Economy*, Vol. 40, p.577-616.

H. Hotelling (1935) "Demand Functions with Limited Budgets", *Econometrica*, Vol. 3, p.66-78.

W. Jaffé (1964) "New Light on an Old Quarrel: Barone's unpublished review of Wickseed's `Essay on the Coordination of the Laws of Distribution" and Related Documents", *Cahiers Vilfredo Pareto*, Vol. 3, p.61-102. As reprinted in D. Walker, 1983, editor, *William Jaffé's Essays on Walras*. Cambridge, UK: Cambridge University Press.

L. Johansen (1959) "Substitution versus Fixed Production Coefficients in the Theory of Growth: A synthesis", *Econometrica*, Vol. 27, p.157-76.

W.E. Johnson (1913) "The Pure Theory of Utility Curves", *Economic Journal*, Vol. 23, p.483-513.

R.F. Kahn (1933) "Notes on the Elasticity of Substitution", *Review of Economic Studies*, Vol. 1.

N. Kaldor (1934) "The Equilibrium of the Firm", *Economic Journal*, Vol. 44.

N. Kaldor (1955) "Alternative Theories of Distribution", *Review of Economic Studies*, Vol. 23 (2), p.83-100.

N. Kaldor (1961) "Capital Accumulation and Economic Growth", in F. Lutz and D. Hague, editors, *The Theory of Capital*, London: Macmillan.

N. Kaldor (1966) *Causes of the Slow Rate of Economic Growth in the United Kingdom*. Cambridge, UK. As reprinted in Kaldor, 1978, *Further Essays on Economic Theory*. New York: Holmes and Meier.

J.M. Keynes (1939) "Relative Movements of Real Wages and Output", *Economic Journal*, Vol. 49, p.34-9.

Tjalling C. Koopmans (1951) "Analysis of Production as an Efficient Combination of Activities", in Koopmans, editor, *Activity Analysis of Production and Allocation: Proceedings of a conference*. New York: Wiley.

F.H. Knight (1921) *Risk, Uncertainty and Profit*. 1933 reprint, London: L.S.E.

F. H. Knight (1928) "A Suggestion for Simplifying the Statement of the General Theory of Price", *Journal of Political Economy*, Vol. 36 (3), p.353-70.

H.D. Kurz and N. Salvadori (1995) *Theory of Production: A long-period analysis*. Cambridge, UK: Cambridge University Press.

W. Leontief (1941) *The Structure of American Economy, 1919-1939: An empirical application of equilibrium analysis*. 1951 edition, New York: Oxford University Press.

A.P. Lerner (1933) "Notes on the Elasticities of Substitution", *Review of Economic Studies*, Vol. 1, p.39-44; p.68-71.

A.P. Lerner (1944) *The Economics of Control: Principles of welfare economics*. New York: Macmillan.

M. Longfield (1834) *Lectures on Political Economy*. 1931 reprint, London: London School of Economics.

F. Machlup (1937) "On the Meaning of Marginal Product", in *Explorations in Economics in honor of Frank Taussig*. New York: McGraw-Hill.

F. Machlup (1935) "The Commonsense of the Elasticity of Substitution", *Review of Economic Studies*, Vol. 2, p.202-13.

T.R. Malthus (1815) *Inquiry into the Nature and Progress of Rent*.

L. Makowski and J.M. Ostroy (1992) "The Existence of Perfectly Competitive Equilibrium *à la* Wicksteed", in Dasgupta, Gale, Hart and Maskin, editors, *Economic Analysis of Markets and Games: Essays in honor of Frank Hahn*. Cambridge, Mass: M.I.T. Press.

H. von Mangoldt (1863) *Principles of Political Economy*.

A. Marshall (1890) *Principles of Economics: An introductory volume*. 1990 reprint of 1920 edition, Philadelphia: Porcupine.

D. McFadden (1963) "Constant Elasticity of Substitution Production Functions", *Review of Economic Studies*, Vol. 30, p.73-83.

D. McFadden (1966) "Cost, Revenue and Profit Functions: A cursory review", Working paper No. 86, University of California, Berkeley. Version reproduced in Fuss and McFadden, 1978.

C. Menger (1871) *Principles of Economics*. 1981 edition of 1971 translation, New York: New York University Press.

K. Menger (1954) "The Logic of Laws of Return: A study in meta-economics", in O. Morgenstern, editor, *Economic Activity Analysis*. Princeton, NJ: Princeton University Press.

M. Morishima (1967) "A Few Suggestions on the Theory of Elasticity" (in Japanese), *Keizai Hyoron* (*Economic Review*), Vol. 16, p.149-50.

M.I. Nadiri (1982) "Producers Theory", in K.J. Arrow and M.D. Intriligator, editors, *Handbook of Mathematical Economics, Vol. II*. Amsterdam: North-Holland.

H. Neisser (1940) "A Note on the Pareto's Theory of Production", *Econometrica*, Vol. 8, p.79-88.

M. Nerlove (1965) *Estimation and Identification of Cobb-Douglas Production Functions*. Skokie, Ill: Rand-McNally.

M. Nerlove (1967) "Recent Empirical Studies of the CES and Related Production Functions", in M. Brown, editor, *Theory and Empirical Analysis of Production*. New York: Columbia University Press.

J. von Neumann (1937) "A Model of General Economic Equilibrium", in K. Menger, editor, *Ergebnisse eines mathematischen Kolloquiums, 1935-36*. 1945 Translation, *Review of Economic Studies*, Vol. 13 (1), p.1-9.

J.M. Ostroy (1980) "The No-Surplus Condition as a Characterization of Perfectly Competitive Equilibrium", *Journal of Economic Theory*, Vol. 22 (2), p.183-207.

J.M. Ostroy (1984) "A Reformulation of the Marginal Productivity Principle", *Econometrica*, Vol. 52 (3), p.599-630.

V. Pareto (1896-8) *Cours d'économie politique*. 2 volumes. 1964 edition.

V. Pareto (1902) "Review of Aupetit", *Revue d'Economie Politique*, Vol. 16, p.90-93.

V. Pareto (1906) *Manual of Political Economy*. 1971 translation of 1927 edition, New York: Augustus M. Kelley.

D. Ricardo (1815) *An Essay on the Influence of a Low Price of Corn on the Profits of Stock*. As reprinted in Volume IV of Ricardo, 1951.

D. Ricardo (1817) *The Principles of Political Economy and Taxation*. London: Dent.

L.C. Robbins (1934) "Remarks upon Certain Aspects of the Theory of Costs", *Economic Journal*.

D.H. Robertson (1931) "Wage Grumbles", in Robertson, *Economic Fragments*, London: P.H. King. As reprinted in American Economic Association, *Readings in the Theory of Income Distribution*. New York: Blakiston.

E.A.G. Robinson (1932) *The Structure of Competitive Industry*. New York:

J. Robinson (1933) *The Economics of Imperfect Competition*. 1950 reprint, London: Macmillan.

J. Robinson (1934) "Euler's Theorem and the Problem of Distribution", *Economic Journal*, Vol. 44, p.398-414.

J. Robinson (1956) *The Accumulation of Capital*. 1969 edition, London: Macmillan.

W.E.G. Salter (1960) *Productivity and Technical Change*. 1966 edition, Cambridge, UK: Cambridge University Press.

P.A. Samuelson (1947) *Foundations of Economic Analysis*. 1983 edition. Cambridge, Mass: Harvard University Press.

P.A. Samuelson (1953) "Prices of Factors and Goods in General Equilibrium", *Review of Economic Studies*, Vol. 21, p.1-20.

P.A. Samuelson (1957) "Wages and Interest: A modern dissection of Marxian economic models", *American Economic Review*, Vol.47, p.884-912.

P.A. Samuelson (1962) "Parable and Realism in Capital Theory: The surrogate production function", *Review of Economic Studies*, Vol. 29, p.193-206.

P.A. Samuelson (1979) "Paul Douglas's Measurement of Production Functions and Marginal Productivities", *Journal of Political Economy*, Vol. 87 (5), p.923-39.

E. Schneider (1934) *Theorie de Produktion*. Wien: Springer.

H. Schultz (1929) "Marginal Productivity and the General Pricing Process", *Journal of Political Economy*, Vol. 37 (5), p.505-51.

G.B. Shaw (1928) *The Intelligent Woman's Guide to Socialism and Capitalism*. Garden City, NY: Garden City Publishing Co.

R.W. Shephard (1953) *Theory of Cost and Production Functions*. 1970 edition, Princeton, NJ: Princeton University Press.

A. Smith (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*. 1981 two-volume reprint of 1976 Oxford edition, Indianapolis, Indiana: Liberty Classics.

R.M. Solow (1958) "A Skeptical Note on the Constancy of Relative Shares", *American Economic Review*, Vol. 48, p.618-31.

R.M. Solow (1962) "Substitution and Fixed Proportions in the Theory of Capital", *Review of Economic Studies*, Vol. 29, p.207-18.

P. Sraffa (1925) "Sulle Relazioni fra Costo e Quantita Prodotta", *Annali di Economia*, Vol. 2 (1), p.277-328.

P. Sraffa (1926) "The Laws of Returns under Competitive Conditions", *Economic Journal*, Vol. 36, p.535-50.

G.J. Stigler (1939) "Production and Distribution in the Short Run", *Journal of Political Economy*, Vol. 48, p.305-27.

G.J. Stigler (1941) *Production and Distribution Theories: The formative period*. 1968 reprint, New York: Agathon.

J.H. von Thünen (1826) *The Isolated State*.

R. Torrens (1815) *Essay on the External Corn Trade*.

H. Uzawa (1963) "Production Functions with Constant Elasticities of Substitution", *Review of Economic Studies*, Vol. 29, p.291-9.

H. Uzawa (1964) "Duality Principles in the Theory of Cost and Production", *International Economic Papers*, Vol. 5, p.216-20.

J. Viner (1932) "Cost Curves and Supply Curves", *Zeitschfirt für Nationalökonomie*, Vol. 3, p.23-46. As reprinted in Stigler and Boulding, 1952, editors, *A.E.A. Readings in Price Theory*. Chicago: Irwin.

F.A. Walker (1876) *The Wages Question*. 1881 edition, New York: Holt.

F.A. Walker (1891) "The Doctrine of Rent and the Residual Claimant Theory of Wages", *Quarterly Journal of Economics*, Vol. 5, p.433-4.

L. Walras (1874) *Elements of Pure Economics: Or the theory of social wealth*. 1954 translation of 1926 edition, Homewood, Ill.: Richard Irwin.

L. Walras (1896) "Note on the Refutation of the English Law of Rent by Mr. Wicksteed", Appendix III to third edition of Walras, 1874.

E. West (1815) *Essay on the Application of Capital to Land*.

K. Wicksell (1893) *Value, Capital and Rent*. 1970 reprint of 1954 edition, New York: Augustus M. Kelley.

K. Wicksell (1900) "Om gränsproduktiviteten sansom grundval för den nationalekonmiska fördelningen" ("Marginal Productivity as the Basis for Economic Distribution"), *Eknomisk Tidskrift*, Vol. 2, p.305-37, as translated in K. Wicksell, 1959.

K. Wicksell (1901) *Lectures on Political Economy, Vol. 1*. 1967 reprint of 1934 edition, New York: Augustus M. Kelley.

K. Wicksell (1902) "Till fördelningsproblemat" ("On the Problem of Distribution"), *Ekonomisk Tidskrift*, Vol. 4, p.424-33, as translated in K. Wicksell, 1959.

K. Wicksell (1916) "Den kritiska punkten i lagen för jordbrukets aftagande produktivitet", *Ekonomisk Tidskrift*, Vol. 18, p.265-92.

K. Wicksell (1923) "Real Capital and Interest", *Ekonomisk Tidskrift*, Vol. 25, p.145-80, as reprinted in Wicksell, 1901.

K. Wicksell (1959) *Selected Papers on Economic Theory*, E. Lindhahl editor, R.S. Stedman translation, London: Allen & Unwin.

P.H. Wicksteed (1894) *Essay on the Co-Ordination of the Laws of Distribution*. 1932 edition, London: L.S.E.

P.H. Wicksteed (1910) *The Common Sense of Political Economy*. 1933 edition, London: Routledge and Kegan Paul.

P.H. Wicksteed (1914) "The Scope and Method of Political Economy in the Light of the "Marginal" Theory of Value and Distribution", *Economic Journal*, Vol. 24 (1), p.1-23.

F. von Wieser (1876) "Über das Verhältnis der Kosten zum Wert" ("On the Relation of Cost to Value"); reprinted in Wieser, *Gesammelte Abhandlungen*, p.377-404.

F. von Wieser (1884) *Über den Ursprung und die Hauptgesetze des wirtschaftlichen Werthes*.

F. von Wieser (1889) *Natural Value*. 1971 reprint of 1893 translation, New York: Augustus M. Kelley.

A.A. Young (1928) "Increasing Returns and Economic Progress", *Economic Journal*, Vol. 28.

References

(1) Paul A. Samuelson, *Foundations of Economic Analysis* (Cambridge, Mass.: Harvard University Press, 1947) [Internet], available at http://cepa.newschool.edu/het/essays/product/prodfunc.htm