The economic market outcomes and income distribution when capital is not homogeneous

Limits of technology

Ahmet Özçam
Department of Economics, Yeditepe University, Istanbul, Turkey

Received 25 December 2018
Revised 6 January 2019
Accepted 10 January 2019

Abstract

Purpose – An aggregate production function has been used in macroeconomic analysis for a long time, even though it seems that it is conceptually confusing and problematic. The purpose of this paper is to argue that the measurement problem related to the heterogenous capital input that exists in macroeconomics is also relevant to microeconomic market situations.

Design/methodology/approach – The author constructed a microeconomic market model to address both the problems of the measurement of the physical capital and of substitutability between labor and capital in the short run using two types of technologies: labor neutral and labor reducing. The author proposed that labor and physical capital inputs are complementary in the short run and can become substitutes only in the long run when the technology advances.

Findings – The author found that even if the technology improves at a fast rate over time, there are then diminishing returns of profits to technology and an upper limit to profits. Moreover, the author showed that under the labor-reducing technology, labor class earns more initially as technology improves, but their incomes start declining after some threshold level of passage of time.

Originality/value – The author cautioned the applied researcher that the estimated labor and capital coefficients of generalized Cobb–Douglas and constant elasticity of substitution of types of production functions could not be interpreted as partial elasticities of labor and capital if in reality the data come from fixed-proportions types of processes.

Keywords High technology, Diminishing rate of profits to technology, Fixed-proportions production technology, Speed as an economic resource, Cobb–Douglas production functions

Paper type Research paper

1. Introduction

There exists a vast economics literature on the aggregation of production functions. An aggregate production function, like $GDP = F(K, L)$, links aggregate inputs into aggregate output and is believed to describe the technology at the aggregate level. It is typically argued that the production function in the theoretical macro model is a microeconomic production function. However, the empirical evidence and examples are from macroeconomics. Such a construct has been used in macroeconomic analysis for a long time, but it seems that it is conceptually confusing and problematic, since the technical conditions which specify a well-behaved aggregate production function that can be derived from micro production functions are so stringent that it is difficult to believe that in applied work one can have those requirements met. Moreover, various attempts by many authors to specify those conditions have each had their own set of assumptions (Leontief, 1947; Gorman, 1953; Fisher, 1969).

Journal of Industry-University Collaboration
Vol. 1 No. 1, 2019
pp. 38-56
Emerald Publishing Limited
2631-357X
DOI 10.1108/JIUC-03-2019-004

With the rise of the new neoclassical endogenous growth literature in the 1980s, a curiosity to test the old and new theories of growth and productivity aroused, especially when large data sets became available. However, these new growth models depended heavily on the old neoclassical aggregate macro production function which was thought to be obtained by adding micro production functions. Since new growth models literature did not seem to be bothered with the old aggregation problem, i.e. without proper aggregation, we cannot study the properties of an aggregate production function (Felipe and Fisher, 2003). This is known as the aggregation problem. Temple (1999) argued that the production function is the least satisfactory element of macroeconomics, yet many economists consider this inept construct as a fundamental basis to understanding national outputs and growth rates. So, it seems that the aggregate production function does not have a sound theoretical basis and therefore is deceiving.

Why do economists keep using aggregate production functions in spite of these difficulties? Since these problems are quite inconvenient for an important part of neoclassical macroeconomics, the profession has simply chosen to ignore them and justify its use by resorting to some stereotypes. For example, since aggregate production functions appear to give empirically satisfactory results, why should not they be used? Or, there is no other choice, and aggregate production functions are seen as parables.

In the economics literature, capital has traditionally been thought as a fund of resources or a set of productive factors (heterogeneous machines) specified in physical units. Of these two concepts, whereas fund or money could easily be measured, real capital was not. Robinson (1953–1954) asked the central question that triggered the so-called Cambridge capital controversies: In what unit is "capital" to be measured? Since capital goods were a series of heterogeneous commodities (investment goods), each having a particular type of technical characteristic embedded in them, it was not possible to express the stock of capital goods as a homogeneous physical entity. Therefore, only their values could be aggregated. Such a value aggregate, however, could not be independent of the rate of profit and therefore of income distribution. So, it was difficult to conceive time series of quantities of capital goods to be combined into a single number that may be used as the amount of real capital in the economy as a whole (Felipe and Fisher, 2003).

The aggregate production function includes two types of aggregation. The first has to do with the meaning of this fictitious construct obtained by aggregating the heterogeneous inputs and outputs (different types of labor, capital and output into one labor, one capital and one output). Although labor is not a homogeneous input, it can naturally be measured in man hours of work[1]. The same is true for land: so many acres of land of a given quality. Since these are natural units, the marginal products of land and labor can be defined independently of the equilibrium factor prices. However, the measurement of the capital input is problematic. The statistics of capital used are in dollars and no matter how they are deflated to convert them into constant dollars, they continue being a monetary value (Felipe and Fisher, 2003). The other aggregation type is summing over firms and in certain contexts over industries in a country or the same industry over different nations.

Fisher (1965) remarked that the assumption that technology is embodied in capital, i.e. that it is firm specific, causes difficulties with respect to the existence of a capital aggregate. He argued that under certain technical conditions, capital-augmenting technical differences among firms, i.e. that the embodiment of new technology can be written as the product of the amount of capital and a coefficient, seem to be the only situation where a capital aggregate exists. Under these circumstances, a unit of one type of new capital equipment is the exact duplicate of a fixed number of old capital equipment. The aggregate stock of capital can be constructed with capital measured in efficiency units (Fisher, 1965). This last comment of Fisher that capital being measured in efficiency turns out what we exactly have in mind to do here. In this paper, we will introduce the concept of "speed" as a measurement of the capacity of real physical capital.

In this paper, we will first argue that the measurement problem related to the capital input that exists in macroeconomics is also present in microeconomics. Similar to the situation in macroeconomics, assume, for example, that the capital investments of some firms in a micro market are made at different points in time and that the profits of these firms decline, due to perhaps a leftward shift in the market demand for the product they sell. This is likely to put a downward pressure on the initial purchasing price (cost) of capital. So, this means that we end up with dissimilar dollar values of costs of investment made at different periods of time (perhaps a few years apart from each other) and it virtually becomes impossible to measure the real quantity of physical capital used in the productions of these firms unless a capital good price index is readily available[2]. Notice that in this example we did not even permit the possibility of heterogeneous capital goods yet that the firms may have used over time due to a possible technological advance.

Moreover, we will further argue that there is also a problem with the use of production functions in microeconomic models which easily allow substitution between labor and capital, where in reality the type of technology may possibly be of fixed-proportions and these proportions may vary among firms or the data collected over time and firms may be of such nature that the technology may have significantly changed so that again the assumed easy substitution between capital and labor may not be so relevant. In applied work, it is common practice in the literature to assume a single type of production function (e.g. Cobb–Douglas and constant elasticity of substitution (CES)) for all the different firms in an industry at a point in time or for the same firms over time while the type of technology may very well change in the meantime. In Section 5, in order to assess the econometric implications of such a stringent assumption which assumes smooth substitution between labor and capital, we will fit both generalized Cobb–Douglas and CES production functions to a hypothetical data set consisting of outputs constructed under four different sets of fixed-proportions production functions where the capital/labor ratios differ (the first representing the oldest technology (most labor-intensive) and the fourth representing the newest technology (most capital-intensive)). As expected, when we estimate such data with these two production functions, some data output levels produced under different technologies (perhaps also with dissimilar physical capital units) will be conceived as if they were obtained with inputs substituted. Put it differently, these typical production functions (Cobb–Douglas and CES) will treat the data output points obtained with heterogeneous types of capital amounts as if they all lied on the same isoquant, i.e. as if the firms were able to switch from one fixed-proportions technology to another quickly in the short run. Clearly, such a substitution can only come about in the long run, and not in the short run when fixed-proportions technologies seem to be more appropriate. Interestingly, even though the non-linear estimation procedure will capture the constant returns to scale (CRS) properties of these four different types of technologies properly using the generalized Cobb–Douglas function, it will not be appropriate to interpret the estimated partial output elasticities of inputs the usual way, since these elasticities are 0 with fixed-proportions production functions[3].

In this paper, we will attempt to construct a microeconomic market model (see Section 3) to address both the problems of the measurement of the physical capital and of substitutability between labor and capital as discussed above. In our model, the firms will be replacing their old machinery/equipment with more productive machinery/equipment embodying important technological advances. There seems to exist only a very limited number of cases/examples, if any, that we can enumerate where labor and capital can be substitutes in the short run. For example, if the dish-washing machine breaks down in a restaurant, then of course, an employee in the kitchen can wash the dishes. However, assume that a large pan is missing in the kitchen of a restaurant. Can labor do the work of such a pan? It seems like that capital and labor are most likely to be complements in the short run, and they can become substitutes in the long run only if a technological improvement occurs. Given an appropriate transformation on the capacity of physical capital input (machinery) that we will introduce (speed), we will show how we will be able to compare heterogeneous types of capital as the technological improvement over time and the sense in which the capital input can be substituted for the labor input in the long run. We will distinguish between two types of technologies: labor neutral and labor reducing. Given the fast technological advances of today, it is clear that the substitution is unidirectional in our era: replacement of capital for labor, and not the other way around.

In Section 2, we will briefly explain the pork production as an example of a simple production process and discuss the extent to which the substitutability between capital and labor is possible and when. In Section 3, we will introduce two types of technologies: labor neutral and labor reducing and our economic resource which is "speed." In the same section, we will also present the short-run and the long-run market outcomes under fixed-proportions production functions. In Section 4, we will discuss the output and profit-maximization effects of the new technology. In Section 5, using hypothetical data, we will show the statistical implications of estimating the technical production parameters of Cobb–Douglas and CES types of production function which allow substitution where by contrast, the data are purposely constructed under fixed-proportions production. Finally, Section 6 concludes the paper.

2. An example of a simple production function where labor and capital are thought to be substitutable: Swine production and its interpretation

We will briefly explain the pork production as an example of a simple production process and discuss to which extent the substitutability between capital and labor is possible and when. Early man, by investing today's labor in the creation of crude tools, made tomorrow's labor easier and/or more productive. Swine confinement, the process of substituting capital for labor, began when pastures were first fenced to reduce the labor needed to harvest the pork (Lines, 1986). The production systems used in pork production can be broadly described as:

(1) high capital/low labor – high-investment confinement;

(2) medium capital/medium labor – low-investment confinement; and

(3) low capital/high labor – pasture systems.

In general, the prices of capital and labor affect the combination of quantities of labor and physical capital preferred by entrepreneurs in any industry. However, there may be some other factors as well:

(1) quantity and type (equity or debt) of money capital available;

(2) the cost of capital (interest rate and/or opportunity cost);

(3) the quantity, quality and timing of available labor besides its cost;

(4) animal performance (feed and breeding efficiency, litter sizes and survival rates);

(5) product prices;

(6) diversification; and

(7) risk aversion and risk-bearing ability of the operator.

As far as the low-technology pork production is concerned, the total cost per unit of production may not vary greatly from one system to another among the three types of combinations of inputs considered above. This may then permit operators, each with a different set of capital, labor and management resources, to compete effectively in

Swine production. The young entrepreneur, generally long on labor and short on money capital, can profitably build a small unit and reduce risk. At the same time, a highly skilled well-financed operator can profitably construct large specialized risky enterprises. The substitutability of capital and labor in the Swine industry may continue to permit the profitable co-existence of diverse production systems (Lines, 1986). However, this almost equality of profits that may have persisted in the example of pork production needs not be the case if a new technology substantially reducing the cost of production is introduced.

The crucial issue here is, of course, the substitutability of capital and labor. We will argue, in this paper, that such an opportunity is only possible in the long run and is not readily available in the short run. Even in the technically relatively simple production setting of pork processing, we cannot assume that entrepreneurs can switch from a combination of a specific type of physical capital and labor to another combination of a different type of machinery and labor in the short run, for example, a quick switch from pastures systems to high-investment confinement. This impossibility which constraints the entrepreneurs is even more pronounced in the situation of today's high-technology industrial production, where the development of automated assembly operations and the increased use of industrial robots have tremendously increased the productive efficiency, but with tremendous efforts of R&D, ingenuity and at a substantial cost of investment which can only be achieved given enough time.

Even if we allow enough time, there is additionally the difficulty of comparing the technologically different quantities (units) of physical capital. Many authors in microeconomics assume that these intrinsically heterogeneous quantities can all be shown on the vertical axis of a typical labor-capital diagram. The quick reachability from one input combination of labor and capital to another combination allows one to derive the well-known smooth downward sloping and convex isoquants. However, in these combinations of labor and capital, first of all, the capital is not homogeneous. Moreover, a different system (technology) may require labor with different types of skills/education/training, and therefore we cannot treat labor homogeneous either. This is simply a substitution of one type of labor with another type. For example, the younger generations will be more apt to learn and understand the new technology. So, it seems that we are substituting non-homogeneous capital units and non-homogeneous amounts of labor within themselves, and not labor with capital. Nevertheless, it is true that in general when technologically more advanced equipment is used, there is need for less labor in the long run.

The price of capital input can of course be measured in terms of monetary expenditures made on machinery per hour. Note, however, that a change in the price of raw material which is essential to machinery manufacturing may affect the capital expenditures. For example, in the case of pork production explained above, a rise in the price of fencing may render low-capital pasture systems more expensive, and hence make it look like more capital-intensive than medium-capital investment confinement. This is an example of a problem associated with measuring physical capital input in monetary terms. Moreover, aggregating such monetary values over time may not be so meaningful.

As an example of capital measured in monetary terms, Arrow *et al.* (1961), in a sample of 19 countries covering 24 manufacturing industries, estimated the rates of return on capital from balance sheets of different industries. Capital was measured by net fixed assets (including land) plus cash and working capital. All financial investments were excluded. Total returns to capital were taken to be equal to gross profit from operations minus depreciation. Again, here we have the problem associated with capital measured in monetary terms.

We now start discussing our market model of fixed-proportions productions with "speed" as a scarce economic resource.

3. A market model of fixed-proportions production functions with speed as an economic resource

In this paper, we distinguish between two types of technologies: labor neutral and labor reducing.

3.1 The labor-neutral and labor-reducing technologies

We are interested in the situation where the firms replace their old machinery/equipment with more productive machinery/equipment embodying important technological advances. For example in standard assembly lines, work assignments, numbers of machines, and production rates are programmed so that all operations along the line are compatible. Automated assembly lines, on the other hand, consist entirely of machines run with little or no human supervision. In such continuous-process industries as petroleum refining and chemical manufacture and in many modern automobile-engine plants, assembly lines are completely mechanized and consist almost entirely of automatic, self-regulating equipment. Some products, however, are still assembled by hand because many component parts are not easily handled by machines. Nevertheless, the development of versatile machinery and the increased use of industrial robots and artificial intelligence have improved the efficiency of fully automated assembly operations. Expensive and somewhat inflexible, automatic assembly machines are economical only if they produce a high level of output within a given period of time (www.britannica.com/technology/assembly-line). To be successful, a new technology must not only deliver higher speed of production, but also do so at an acceptable cost (price of machinery).

In the recent past, there has been an extraordinary decrease in prices of personal computers and equally impressive advance in the computing power and other tasks performed with the typical personal computer. The demand for workers depended on the type of production function. It has increased for some types of workers, while it has decreased for others. The demand for software computer programmers and computer-assembly workers has soared rapidly. However, the demand for office workers has been mixed (McConnell and Brue, 1999). In some offices, the higher speed of the personal computers has increased the demand for computer engineers, since in this case, the computers and the engineers are complementary inputs, whereas in some other cases, the firms substituted computers for labor and decreased their demand for labor.

In this paper, we distinguish between two types of technologies: labor neutral and labor reducing. In Table I, we display a numerical example of a labor-neutral technology. Technology A is the firm's current process of production. It uses 1 man-hour of labor and 4 machine-hours of capital per unit of output. Assume that a new technology of production, B, is introduced resulting in productive efficiency which uses the same man-hours of labor (hence labor neutral) but uses 3 machine-hours of capital per unit of output (process innovation rather than product innovation). Technology C captures even more opportunities in that it still employs 1 man-hour but only 2 machine-hours. We assume that the capability of the physical capital input determines the duration of the production of one unit of output and the firm adjusts its demand for labor accordingly. In other words, machinery is the more important factor of production. For example, with the Technology A, the firm purchases four machines and uses 16 machine-hours and

	Labor requirement	Capital requirement	Speed of capital
Technology A	1 man-hour	4 machine-hours	0.25
Technology B	1 man-hour	3 machine-hours	0.333
Technology C	1 man-hour	2 machine-hours	0.5

Table I. Labor-neutral technologies (to produce 1 unit of output)

4 man-hours of labor to produce 4 units of output at the end of the fourth hour, so that no input stays idle.

In Table II, we show an example of a labor-reducing technology. Current technology is again A and uses 1 man-hour of labor and 4 machine-hours of capital per unit of output, as shown in Table I. However, the new technology of production D now requires only 0.95 man-hours, while the physical capital requirement stays the same as before (3 machine-hours per output). Technology E employs even less of labor, only 0.90 man-hours.

3.2 The short run vs the long run and speed as an economic resource

In standard microeconomics textbooks, the production function is drawn where the amount of labor is put on the horizontal axis and the quantity of physical capital of different technologies on the vertical axis as if the capital were a homogeneous input. Perloff (2012, p. 189), for example, discussed the isoquants of semiconductor circuits (or "chips") in order to give a real-world example. He discussed three different types of technology:

(1) aligners: least capital-intensive process;

(2) steppers: more capital-intensive process; and

(3) steppers with wafer-handling equipment: most capital-intensive process.

In all of these three cases, the output is supposed to be 25 ten-layer chips. Capital lies on the vertical axis and labor on the horizontal axis, as usual. Perloff then proceeds to obtaining downward sloping piecewise-linear isoquants by combining different combinations of fixed-proportions of labor and capital. If more combinations are considered, then in the limit, the familiar smooth downward sloping convex isoquants representing perfectly substitutability among inputs as asserted in the theory of microeconomics are pretended to prevail. However, "amounts of capital" do not have a common unit of measurement in different technological processes and hence cannot be shown on the same axis. Units of apples and oranges cannot be compared. Moreover, an attempt to draw an isocost line would also be not well grounded since technological units unlike in form cannot be combined with a single price.

We now introduce our concept of "speed" which can be used to compare technologically different units of physical capital:

Definition 1. The "speed of physical capital" is the inverse of the machine-hours to produce 1 unit of output.

In other words, it is the number of output a machine can produce in a unit of time (e.g. an hour). The speeds of Technologies A–E are shown in the fourth column of Tables I and II. Productive efficiency with respect to process innovation in economics is defined as the ability to produce the same amount of output using fewer amounts of inputs (or to produce more with the same quantity of inputs). Unfortunately, this definition of efficiency does not seem to be appropriate in the case of a technological advance. When some technological advance occurs so that efficiency occurs, the nature of some of the inputs is expected to change (i.e. a more automated process), so that the levels of inputs before and after the technological change cannot be compared. This is clearly a contradiction. With the

Table II. Labor-reducing technologies (to produce 1 unit of output)

	Labor requirement	Capital requirement	Speed of capital
Technology A	1 man-hour	4 machine-hours	0.25
Technology D	0.95 man-hours	3 machine-hours	0.333
Technology E	0.90 man-hours	2 machine-hours	0.5

introduction of our term "speed of capital," we are defining productive efficiency as the capability of physical capital (including robots and artificial intelligence) to produce more per unit of time (i.e. an hour) and, therefore, both types of physical capital, old and new, are replaced with the same variable (speed).

Based on all the real-world examples discussed above, we will propose the following:

P1. In general, labor and physical capital inputs are complementary in the short run and can become substitutes only in the long run when the technology advances.

In Section 5, we will discuss the econometric consequences of estimating production functions where the applied researcher would assume perfect substitutability between inputs, whereas in reality, we constructed the hypothetical data according to fixed-proportions technology.

3.3 The short-run market equilibrium with fixed-proportions production function

Many characteristics of the oligopolistic market structure help technological advance. First of all, oligopolists realize economic profits, some of which may be retained. This undistributed profit may serve as a major source of funding their R&D expenditures. Second, the existence of possible entry barriers gives the oligopolists an opportunity to continue any economic profit they receive from innovation. Third, large sales volumes allow them to have smaller R&D expenditures per unit of output. Finally, the broad nature of R&D activity within oligopolistic firms helps them counterbalance the inevitable unsuccessful innovative attempts with the successful ones. Therefore, in general, the oligopolists have the means and incentive to innovate even though introducing costly new technology may render some of its current equipment obsolete. On the other hand, a pure monopolist protected with high entry barriers may tend to breed complacency (McConnell and Brue, 1999).

We start with the Cournot's model with N firms producing a homogeneous good. Assuming that the market demand curve is linear with $p=a-bQ$ (where p is the market price, Q is market output and a and b are some positive real numbers). We further assume that each firm i ($i=1, 2, \dots, N$) is using a production process characterized with fixed-proportions of labor, capital and raw material/intermediary goods in the short run based on our *P1*. Concentrating on the production of 1 unit of output, the constant marginal cost of the ith firm is:

$$mc_i = w \times \alpha + p_K \times \frac{1}{S} + p_{RAW}, \tag{1}$$

where S is the speed of physical capital, α and β are man-hours and machine-hours required to produce 1 unit of output, respectively, and $S=(1/\beta)$. w, p_K and p_{RAW} are the wage rate, rental price of capital and cost of raw material (or intermediary inputs) all used in producing 1 unit of output, respectively. Notice that the prices of inputs are defined with respect to the production of 1 unit of output and not per period of time. This seems to be necessary since all prices must have a common dimension in the calculation of any type of cost function. Even though the labor and capital are perhaps rented, there is a fixed amount of cost of raw material (or intermediary inputs) for unit of output that is produced. For example, if the capital is rented for $3,200 a week, and the production setting of the plant is such that 200 units are planned to be produced per week, then the price of capital is $16 per unit of output[4]:

P2. It seems that it makes more sense to calculate the costs of inputs per unit of output rather than per period of time. After all, in the case when all inputs are complementary in the short run and some inputs stay idle, this does not help any further production.

Each firm maximizes its profits:

$$\Pi_i = p \times q_i - average\_cost_i \times q_i = p \times q_i - mc_i \times q_i, \tag{2}$$

since the marginal cost is the same thing as the average cost under fixed-proportions production. Equality of marginal revenue and marginal cost gives:

$$\frac{d\Pi_i}{dq_i} = p + q_i \times \frac{dp}{dQ}\left(\frac{dQ}{dq_i}\right) = mc_i, \tag{3}$$

where $(dQ/dq_i) = 1$, due to Cournot type firms' expectations. Also:

$$a - b \times (q_1 + q_2 + \cdots + q_N) - b \times q_i - mc_i = 0. \tag{4}$$

Summing over all N symmetric firms having the same marginal cost ($mc_i = mc$), we have:

$$Na - NbQ - bQ - Nmc = 0, \tag{5}$$

and the usual equilibria market price and market output are:

$$Q_e = \frac{N \times (a - mc)}{b \times (N+1)} \text{ and } p_e = \frac{(a + N \times mc)}{(N+1)}. \tag{6}$$

3.3.1 Numerical example 1 (short run). Assume that $w = \$10$, $p_K = \$16$, $p_{RAW} = \$3$, $N = 2$, $a = 120$, $b = 1$ and $\alpha = 1$ man-hour, $\beta = 4$ machine-hours (as in Technology A in Table I) so that $S = 0.25$ outputs and $mc = \$10 \times 1 + \$16/0.25 + \$3 = \$77$. We obtain $Q_e = 28.67$ units of output and $p_e = \$91.33$ using Equation (6). The market equilibrium level of profits is $\Pi_e^{Market} = (91.33 - 77) \times 28.67 = \$410.89$. The demand for labor derived from the market demand curve is $\alpha \times Q_e = 28.67$ man-hours of labor and total wages paid to the labor in the industry are \$286.66 (28.67 units $\times$\$10).

Note that the equilibrium price decreases and the equilibrium quantity of market output increases with the number of firms, N, in the industry the usual way. The Lerner Index (LI) of monopoly power of each of the two firms is as follows:

$$LI = \frac{(p - mc)}{p} = \frac{1}{-\epsilon^f} = \frac{1}{(-(dQ \times p/dp \times Q) \times 2)} = 0.1569, \tag{7}$$

where ϵ^f is the price elasticity of demand facing each of the two firms.

We now start discussing the long-run equilibria of this market when technology improves over time.

3.4 A long-run model where technology may induce substitution between labor and technologically improved physical capital

3.4.1 Long run: labor-neutral technology. In the same industry, we now assume that the speed of technology of production grows at a rate given by $g_1 = \exp(time/c)$ so that the speed of newer technology becomes $S \times g_1$ with respect to that of the initial Technology A which is S. The price of physical capital increases at a rate of $g_2 = \exp(time/d)$, where we assume that $c < d$, and therefore $g_1 > g_2$ for each time period, i.e. the growth of speed of technology is greater than the growth of the price of physical capital. Otherwise, there would be no motive to use the new technology. The marginal cost under labor-neutral technology mc_{LN} then becomes:

$$mc_{LN} = w \times \alpha + \left(\frac{g_2 \times p_K}{g_1 \times S}\right) + p_{RAW}. \tag{8}$$

Clearly, then mc_{LN} decreases as new technologies are adopted[5]:

P3. Under the labor-neutral technology, even if the technology improves at a fast rate over time, there is a lower limit for the marginal cost of producing and therefore a lower limit for the price of the product and an upper limit for the market quantity of output.

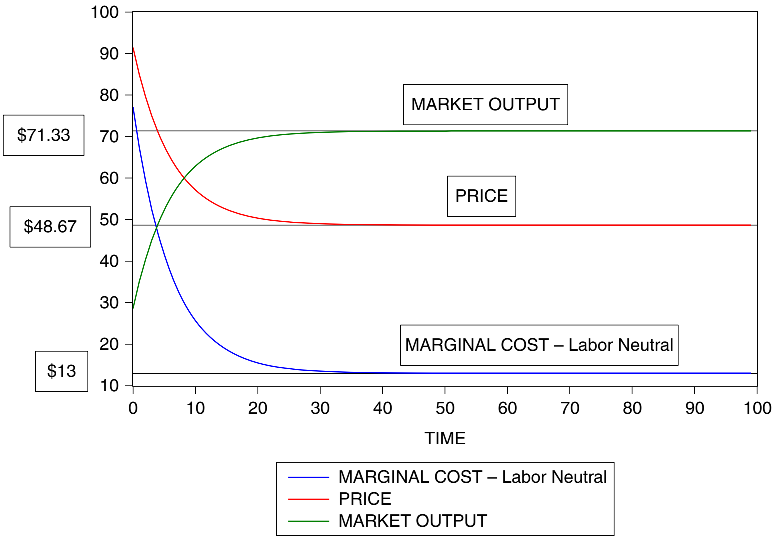
Proof. The lower limit for mc_{LN} is $w\times\alpha+p_{RAW}$, since the limit of $(g_2\times P_K/g_1\times S)$ is 0 as time goes to infinity as long as $g_1 > g_2$. When $w\times\alpha+p_{RAW}$ is substituted for the marginal cost in Equation (6) for the equilibrium quantity, Q_e, and the equilibrium price, p_e, we see immediately that the market equilibrium output has an upper bound and the market equilibrium price has a lower bound. ■

3.4.2 Numerical example 2 (long run, labor-neutral technology). Assume that the values of w, p_K, p_{RAW}, N, a, b, α and S are the same as in the numerical example 1. Assume further that $c=3.48$. Then, the speed of new technology at time Period 1 is $S\times g_1=0.25\times\exp(1/3.48)$ $=0.25\times1.332=0.333$ that corresponds to Technology B in Table I. We take $d=8$.

We now suppose that the technology improves steadily over time as discussed before. In Figure 1, we see that the marginal cost under labor-neutral technology reaches a lower bound given by \$13 (\$10× 1+\$3) in this case determined by the prices of labor and raw material inputs. As a result of that, the price of the product approaches a lower bound of \$48.67 and the market quantity of output reaches an upper bound 71.33 units, as time passes (again by Equation (6)).

In Figure 2, we observe that both market profits and wages benefit from the technology which causes the market output to expand, but disproportionately. In this numerical example, while total market profits increase from \$411 to \$2,544, total market wages increase from \$287 to only \$713. Therefore, both social classes become more affluent under labor-neutral technology, but the capital owners more so, in the oligopolistic market situation:

P4. Under the labor-neutral technology, even if the technology improves at a fast rate over time, there are positive diminishing returns of profits to technology and an



Notes: $N=2$; $a=120$; $b=1$; $\alpha=1$ man-hour; $\beta=4$ machine-hours; $w=\$10$; $p_K=\$16$; $p_{RAW}=\$3$; $c=3.488$; $d=8$

Figure 1. The market outcomes in the long run: labor-neutral technology

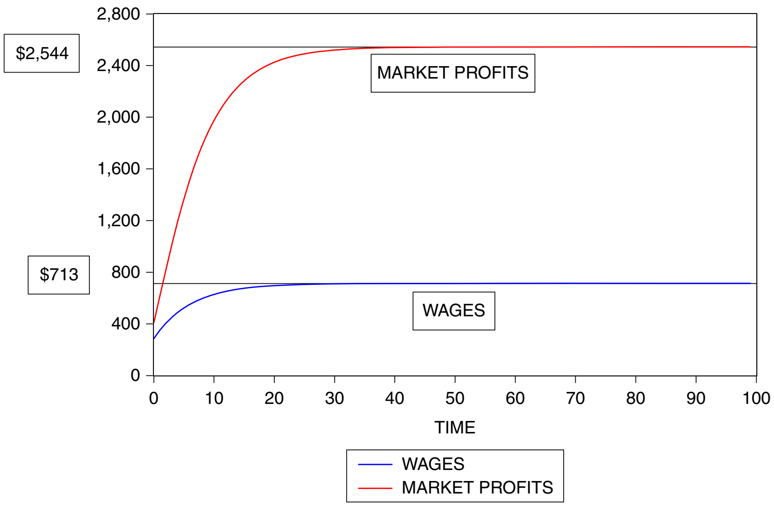


Figure 2.
The income distribution of labor and profits in the long run: labor-neutral technology

upper limit to profits determined by other characteristics/parameters of the market, i.e. the market demand curve and prices of labor, raw material inputs.

Proof. Since there are lower limits to the marginal cost and the price of the product and an upper limit to the market output, as given by *P2*, there is also an upper limit to profits, $\Pi_e^{Market} = (p_e - mc_{LN}) \times Q_e$, i.e. profits do not increase indefinitely. The returns to profits and wages are positive but diminishing since at the end, they must reach their respective upper bounds. ■

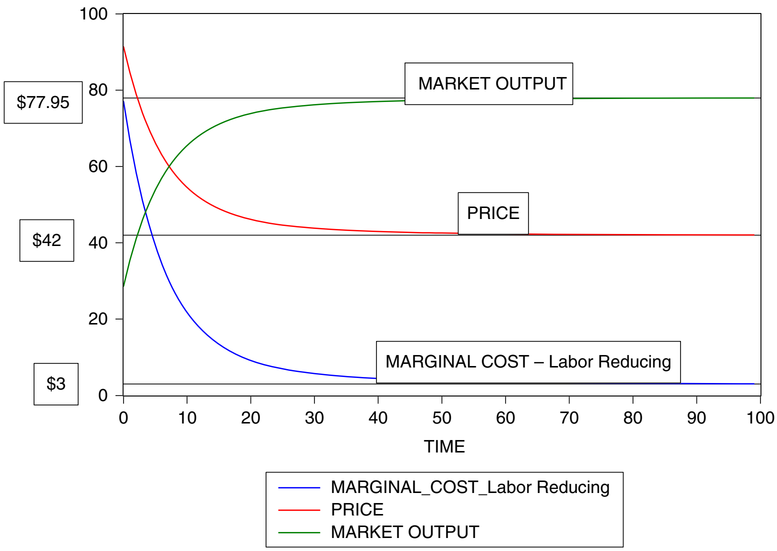
3.4.3 Long run: labor-reducing technology. We now assume that everything is the same in the industry as before except that the new technology needs less labor to work with advanced physical capital equipment: hence, labor-reducing technology. This situation allows for a substitution between labor and capital in the long run. To be more precise, it is substitution between labor and capital which changes due to technology; hence, we have a situation where physical capital is heterogeneous. The demand for labor is $\alpha/g_3 = \alpha/\exp(time/l)$, where l is constant. The marginal cost under labor-reducing technology, mc_{LR}, then becomes:

$$mc_{LN} = w \times \left(\frac{\alpha}{g_3}\right) + \left(\frac{g_2 \times p_K}{g_1 \times S}\right) + p_{RAW}. \tag{9}$$

3.4.4 Numerical example 3 (long run, labor-reducing technology). Assume that the values of $w, p_K, p_{RAW}, N, a, b, \alpha, S, c$, and d are the same as in the numerical example 2. Assume further that $l = 20$. Then, the demand for labor at time Period 1 is $\alpha/g_3 = \alpha/\exp(time/l) = (1/\exp(1/20)) = 0.95$ man-hours and it corresponds to Technology D in Table II.

We now suppose that the labor-reducing technology improves steadily over time. In Figure 3, we see now that the marginal cost under labor-reducing technology reaches a bound which is even lower than in Example 2 and it is given by \$3 which is the cost of only the raw material. Similarly, the price of the product and the market quantity of output approach upper bounds of \$42 and 77.95 units, respectively, as time passes.

In Figure 4, we observe that the markets profits now increase up to \$3,038 and are higher than \$2,544 of under the labor-neutral technology. The interesting result is the dismal outcome for wages. Even though wages first increase from \$287 up to \$421, they start declining steadily afterwards all the way down to \$5.52 We can say that in the rising portion



Notes: N=2; a=120; b=1; α=1 man-hour; β=4 machine-hours; w=\$10; p_K=\$16; p_{RAW}=\$3; c=3.488; d=8; l=3.488×15=52.32

Figure 3.
The market outcomes in the long run: labor-reducing technology

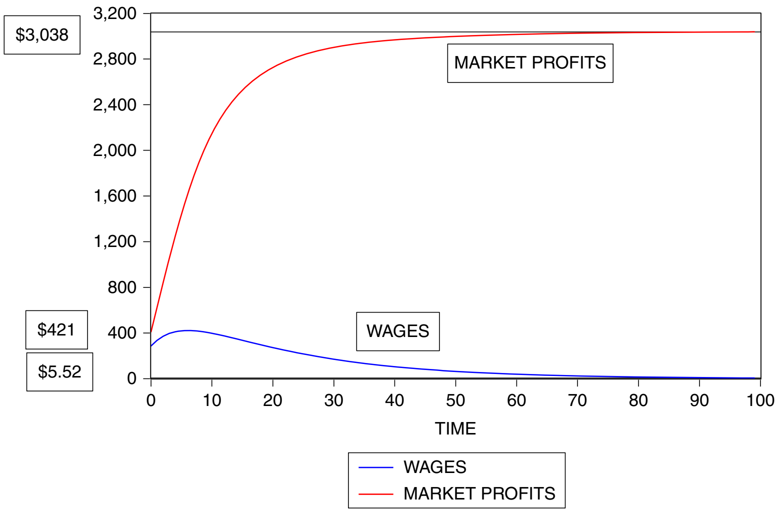


Figure 4.
The income distribution of labor and profits in the long run: labor-reducing technology

of wages, the rate of increase in the market output and therefore an increased demand for labor dominate the lesser amount of labor required due to the advancing technology. In the declining portion of wages function, the latter factor unfortunately outweighs the former:

P5. Under the labor-reducing technology, labor class earn more initially as technology improves, but their incomes start declining after some threshold level of passage of time.

The wages are equal to the demand for labor times the fixed wage rate, which is \$10 in our numerical example. Therefore, the labor demand function in Figure 5 reflects the same concave

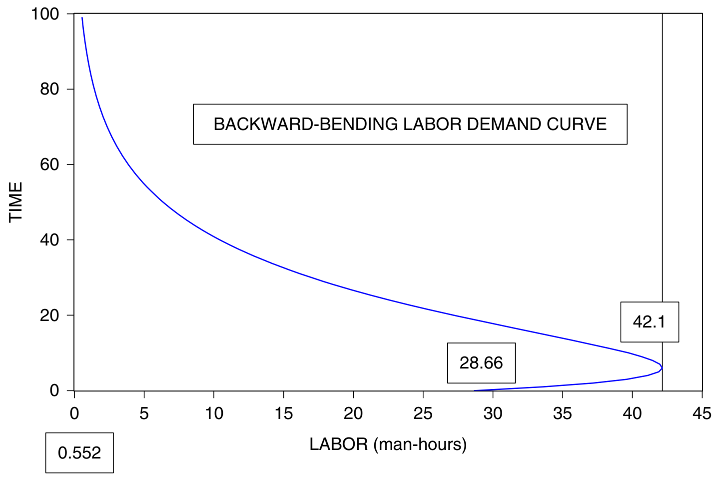


Figure 5.
Backward-bending demand for labor: labor-reducing technology

shape as wages in Figure 4, but this time, the time variable is drawn on the vertical axis. In Figure 5, the demand for labor starts at an initial level of 28.66 man-hours, as before under the initial technology A. However, the labor demand starts backward-bending after it reaches a level of 42.1 man-hours. At the end, the demand for labor is as low as 0.552 man-hours:

P6. Under the labor-reducing technology, the demand for labor increases first but starts decreasing eventually, i.e. the labor demand bends backward.

Proof. the rate of increase in the market output becomes less and less under labor-reducing technology similar to *P3* and therefore initially an increased demand for labor dominates the lesser amount of labor required under the more advanced technologies. However, eventually, the latter factor unfortunately outweighs the former and the labor demand bends backward. ■

We can compare the LIs which show the market powers of firms under these two types of technologies. This is shown in Figure 6. Firms have more market power under the

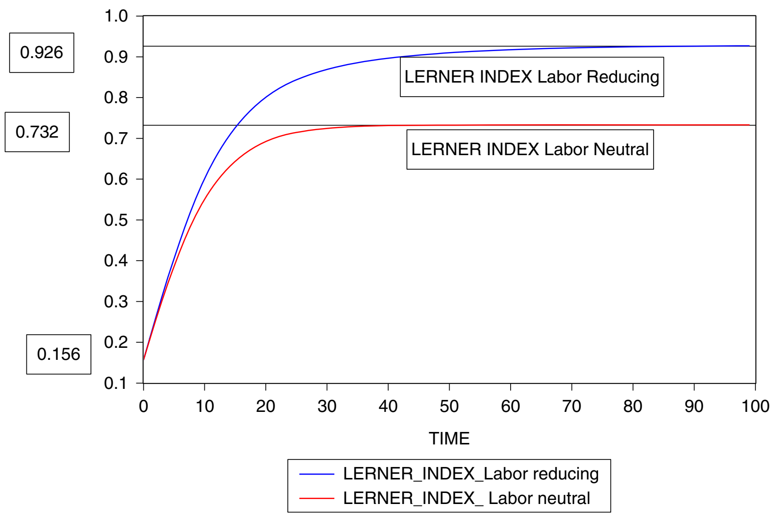


Figure 6.
The Lerner indexes under labor-neutral and labor-reducing technologies

labor-reducing technologies compared to those under labor-neutral technologies. According to the deterministic approach to the market concentration or scale economies hypothesis, it is well known that as new technologies are introduced, long-run average cost curve shift downward and to the right so that some firms unable to adapt themselves exit from the industry in the long run with the consequence that the market concentration increases. In our model, we did not allow for entry or exit in the long run as technology improved. Therefore, the market concentration measured with either, for example, n-firm concentration ratio or Hirschman–Hirfindahl index stays the same in our model. However, we find firms to have more market power in the sense that they can raise their prices more above their marginal costs proportionately:

P7. Firms have more market power under the labor-reducing technologies compared to those under labor-neutral technologies.

4. Output and profit-maximization effects of technology

It is clear that as the technology improves, the market output increases. We can decompose this rise in output into two separate components:

(1) the output effect, which is the extra amount that can be produced with the same amount of monetary expenditures as before, due to technological advance; and

(2) the profit-maximization effect, which prompts even more output to be produced since the marginal cost to be equal to the marginal revenue is required by profit maximization.

Using our Example 2, we found the initial market output under Technology A to be 28.67 units and the market output under Technology B to be 35.04 units and the difference being 6.37 units. The expenditure to produce 28.67 units with the marginal cost of $77 is $2,207.57. When the marginal cost decreases down to $67.44 under Technology B, the same expenditure amounts of $2,207.57, more output to be produced. This is 32.73 units. Then, the output effect is 32.72–28.67 or 4.06 units and the profit-maximization effect is 35.04–32.72 or 2.32 units.

We now turn to the econometric estimation of our hypothetical data with four fixed-proportions production technologies.

5. Econometric estimation of fixed-proportions production functions

In this section, we will estimate the technical parameters of generalized Cobb–Douglas and CES of types of production functions which assume smooth substitution among inputs. In contrast, our hypothetical data set is constructed to represent four different types of fixed-proportions production functions in a sample size of 20, which are given in Table III. It assumed that the data on these four different sets of inputs–output combinations represent either the same five firms over time adopting four progressively more advanced technologies represented by Process 1 (Obs. No. 1–5), Process 2 (Obs. No. 6–10) … and Process 4 (Obs. No. 16–20), processes shown as four fixed input ratios along 4 rays from the origin in a typical labor-capital diagram (like Figure 7), or the same data of 20 observations are assumed to represent four groups of five firms in an industry at a given point in time, each group using its own fixed-proportions processes shown as again four fixed input ratios along four rays from the origin, resulting in five different levels of outputs for each group of five firms. The input ratios (capital/labor) for these four groups of data are taken to be 2 (Obs. No. 1–5), 2.75 (Obs. No. 6–10), 4 (Obs. No. 11–15) and 7.5 (Obs. No. 16–20). The output level of 1 unit (in the fourth column of Table III) is obtained with input combinations given by Observation No. 1, 6, 11 and 16. The remaining output levels (other rows of Column 4)

Obs. No.	Labor (L)	Capital (K)	Output (Q)
1	5	10	1
2	12	24	2.4
3	16	32	3.2
4	20	40	4
5	25	50	5
6	4	11	1
7	9	24.75	2.25
8	11	30.25	2.75
9	16	44	4
10	20	55	5
11	3	12	1
12	7	28	2.33
13	9	36	3
14	13	52	4.33
15	15	60	5
16	2	15	1
17	3	22.5	1.5
18	6	45	3
19	9	67.5	4.5
20	10	75	5

Table III.
Hypothetical data on four fixed-proportions production functions

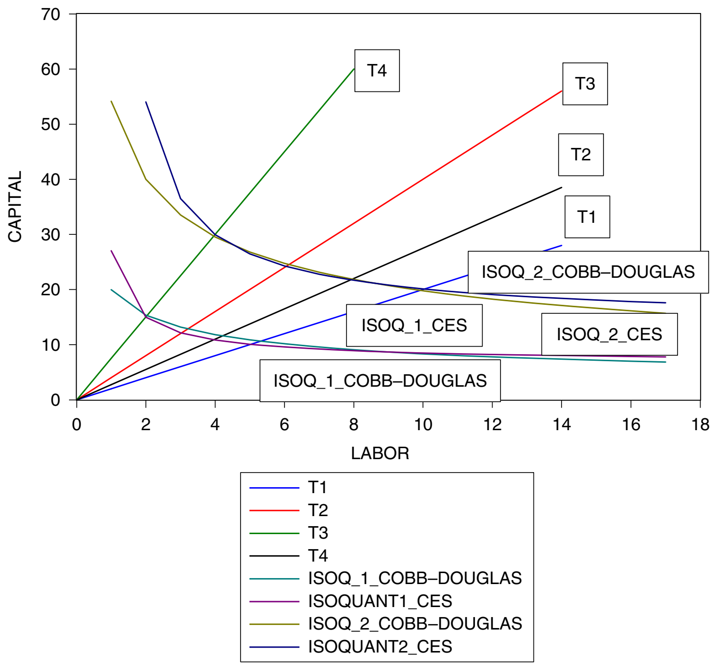


Figure 7.
Fixed-proportions technologies and the estimation of Cobb–Douglas and CES production functions

correspond to the same multiples of the capital and labor amounts keeping capital/labor ratios constant in each group to have a fixed-proportions production function. For example, in the first group (the first five observations), the capital-labor ratio is fixed, $K/L = 10/5 = 24/12 = 32/16 \ldots 50/25 = 2$ and the output levels are 1, 2.4, 3.2, 4 and 5 units.

Hence, whether the data refer to fixed-proportions production functions changing over time, which is a technological advance (case 1 above) or 4 different types of technologies existing concurrently at a point in time (case 2 above), there is clearly no possibility of substitution between labor and capital inputs in the short run.

We also suppose that there is no problem of measurement of physical capital, since our aim is to investigate the econometric consequences of estimating 4 different sets of fixed-proportions of input combinations representing technologically dissimilar processes, using two typical production functions, generalized Cobb–Douglas and CES which do allow substitution among inputs. These two production functions are shown in the following equations:

$$Q = A \times Labor^{\alpha} \times Capital^{\beta} \text{(Generalized Cobb–Douglas)}, \tag{10}$$

$$Q = A \times (\delta \times Capital^{-\rho} + (1-\delta) \times Labor^{-\rho})^{-1/\rho} \text{(CES)}, \tag{11}$$

where α and β are coefficients of labor and capital, respectively, in Equation (10), and δ and ρ are the distribution and substitution parameters, respectively, in Equation (11). A is the scale or efficiency parameter. The estimation results are given in Equations (12) and (13), using non-linear least squares estimation method in Eviews, where all coefficients are highly significant (p-values almost 0):

$$Q = 0.1249 \times Labor^{0.304} \times Capital^{0.695} \text{(Generalized Cobb–Douglas)}, \tag{12}$$

$$R^2 = 0.99,$$

$$Q = 0.1145 \times \left(0.8258 \times Capital^{-0.525} + (1-0.8285) \times Labor^{-0.525}\right)^{-1/0.525} \text{(CES)}, \tag{13}$$

$$R^2 = 0.99.$$

The fits in both equations are almost perfect ($R^2 = 0.99$), since we constructed the data in such a way that the output levels correspond to the same multiples of the capital and labor amounts keeping *K/L* constant, as mentioned before and shown in Table III.

The isoquants derived from these two types of production functions are as follows:

$$Capital = \left(\frac{\overline{Q}}{A} \times Labor^{-\alpha}\right)^{1/\beta} \text{(Generalized Cobb–Douglas)}, \tag{14}$$

$$Capital = \left(\frac{\left(\overline{Q}/A\right)^{-\rho} - (1-\delta) \times Labor^{-\rho}}{\delta}\right)^{-1/\rho} \text{(CES)}. \tag{15}$$

In Figure 7, T1, T2, T3 and T4 are some straight lines displaying the fixed input ratios representing 4 different types of processes discussed above, and the ones having relatively higher capital/labor (*K/L*) ratios presumably representing more advanced technologies. The estimated isoquants corresponding to the output levels of $Q = 1$ and $Q = 2$ are also shown in the same figure.

Our findings in this example are as follows:

(1) The generalized Cobb–Douglas function correctly estimates that there is CRS, since the sum of the coefficients of labor and capital adds up to one ($0.304+0.695=0.999$). This needed not to be the case since we did not impose any restrictions on these elasticity parameters (hence, the name generalized Cobb–Douglas function). So far so good! However, these labor and capital coefficients cannot be interpreted as partial elasticities of labor and capital, since in our example, their marginal products of inputs are precisely zero by the construction of fixed-proportions production functions.

(2) The CES function, which is inherently characterized by CRS, assigns different coefficients to labor and capital from those obtained with the Cobb–Douglas case above in part (1): 0.174 and 0.826, respectively. This result is a consequence of the fact that the elasticities of substitutions of these two different types of functions differ. Whereas the Cobb–Douglas is characterized by unitary elasticity of substitution, the elasticity of substitution of the CES is estimated to be 0.656 ($(1/1+\rho)=1/(1+0.525)$). Again this elasticity of substitution does not make sense in our present context, since we cannot speak of any substitution at all due to the fixed-proportions nature of the data.

(3) As expected, these two well-known production functions conceive various constant output levels (e.g. Isoquants 1 and 2 in Figure 7) produced with dissimilar physical capital-labor ratios presumably representing different processes as if substitution were possible, i.e. as if the firms were able to switch from one technology to another quickly in the short run. Interestingly enough, we find a quite a successful fit as the isoquants for $Q=1$ and $Q=2$ nearly pass through the points of 1 unit and 2 units of outputs indicated by T1, T2, T3 and T4 lines starting from the origin and representing fixed input combinations (see Figure 7).

Our main conclusion is then the following. If, in data, dissimilar production processes are mixed, then the isoquants shown in Figure 7 can, perhaps at best, be interpreted as a substitution between the capital and the labor in the long run. But then, again we face the problem of measurement of capital. Heterogeneous units of capital depicting dissimilar technologies cannot be measured on the same axis. If they are measured anyways in monetary terms (in dollars) and we accept, for example, \$1,000 as a numeraire as it is frequently done in applied work, \$1,000's worth of capital of old technology is not the same thing as \$1,000's worth of capital of new technology. That was our rationale of introducing the concept of "speed" as a unit of measurement for physical capital representing machinery in Section 3.2.

6. Conclusion

An aggregate production function has been used in macroeconomic analysis for a long time, even though it seems that it is conceptually confusing and problematic. In this paper, we argued that the measurement problem related to the heterogenous capital input that exists in macroeconomics is also relevant to microeconomic market situations. Moreover, we further argued that there was also a problem in using in applied work: the typical production functions in microeconomic theory which allowed substitution between labor and capital in a case where the data might represent outputs obtained under different fixed-proportions technologies adopted in the long run. We constructed a microeconomic market model where the firms replaced their old machinery/equipment with more productive machinery/equipment embodying important technological advances. We introduced the term "speed" and tried to show how we have been able to compare heterogeneous types of physical capital as the technological improvement over time. We distinguished between two types of technologies: labor neutral and labor reducing. Given the fast technological advances of today, it is clear that the substitution is unidirectional in our era: replacement of capital for labor, and not the other way around.

We made the following seven propositions:

P1. In general, labor and physical capital inputs are complementary in the short run and can become substitutes only in the long run when the technology advances.

P2. It seems that it makes more sense to calculate the costs of inputs per unit of output rather than per period of time. After all, in the case when all inputs are complementary in the short run and some inputs stay idle, this does not help any further production.

P3. Under the labor-neutral technology, even if the technology improves at a fast rate over time, there is a lower limit for the marginal cost of producing and therefore a lower limit for the price of the product and an upper limit for the market quantity of output.

P4. Under the labor-neutral technology, even if the technology improves at a fast rate over time, there are positive diminishing returns of profits to technology and an upper limit to profits determined by other characteristics/parameters of the market, i.e. the market demand curve and prices of labor, raw material inputs.

P5. Under the labor-reducing technology, labor class earns more initially as technology improves, but their incomes start declining after some threshold level of passage of time.

P6. Under the labor-reducing technology, the demand for labor increases first but starts decreasing eventually, i.e. the labor demand bends backward.

P7. Firms have more market power under the labor-reducing technologies compared to those under labor-neutral technologies.

Finally, we estimated the parameters of generalized Cobb–Douglas and CES of types of production functions using a hypothetical data set representing four different fixed-proportions production functions. We cautioned the applied researcher that these labor and capital coefficients could not be interpreted as partial elasticities of labor and capital, since the marginal products of inputs were 0 under fixed-proportions production functions.

Notes

1. However, there are certainly some instances where different types of labor can substitute each other and therefore represent examples where labor is not homogeneous either. For example, a crafted carpenter who is paid $20 per hour finishes a task in 8 hours using $100 in materials costing $260 for the day of work. It might take an unskilled laborer 20 hours to complete the same job working for $10 per hour and using $150 in materials costing $350 in total. Cost minimizing requires hiring the skilled laborer who does the job faster using less costly material.
2. Some other factors on the supply side may as well affect the initial purchasing price (cost) of capital.
3. On the other hand, the CES types of production functions display CRS by construction.
4. Of course, it is much more difficult to calculate the cost of capital in a production context when the capital is owned.
5. Using the new technology and discarding the old one are economical since their marginal cost will be lower. However, renewing all the machines may take some time and in the short run, the firm may continue to use its old technology machines because of possible financing constraints. If the machines that the firm are employing are rented, rather than owned, then renewing can be accomplished at the end of the lease period and the diffusion of the new technology is expected to be much faster.

References

Arrow, K.J., Chenery, H.B., Minhas, B.S. and Solow, R.M. (1961), "Capital-labor substitution and economic efficiency", *Review of Economics and Statistics*, Vol. 43 No. 3, pp. 225-250.

Felipe, J. and Fisher, F. (2003), "Aggregate production functions: what applied economists should know", *Metroeconomica*, Vol. 54 No. 2, pp. 208-262.

Fisher, F.M. (1965), "Embodied technical change and the existence of an aggregate capital stock", *Review of Economic Studies*, Vol. 37 No. 4, pp. 553-577.

Fisher, F.M. (1969), "The existence of aggregate production functions", *Econometrica*, Vol. 37 No. 4, pp. 553-577.

Gorman, W.M. (1953), "Community preference fields", *Econometrica*, Vol. 21, pp. 63-80.

Leontief, W.W. (1947), "Introduction to a theory of the internal structure of functional relationships", *Econometrica*, Vol. 15 No. 4, pp. 361-373.

Lines, A.E. (1986), "Capital-labor substitution", prepared for 1986 American Pork Congress, St Louis, MO, March 11–13.

McConnell, C. and Brue, S. (1999), *Economics*, Irwin McGraw-Hill.

Perloff, J.M. (2012), *Microeconomics*, Pearson.

Robinson, J. (1953–1954), "The production function and the theory of capital", *Review of Economic Studies*, Vol. 21 No. 2, pp. 81-106.

Temple, J. (1999), "The new growth evidence", *Journal of Economic Literature*, Vol. 37, March, pp. 112-156.

Web reference

Assembly Line, Industrial Engineering, available at: www.britannica.com/technology/assembly-line (accessed August 25, 2018).

Further reading

Chiang, A.C. (2005), *Fundamental Methods of Mathematical Economics*, McGraw-Hill Book.

Corresponding author
Ahmet Özçam can be contacted at: aozcam@yeditepe.edu.tr