Centro di Ricerche
e Documentazione
"Piero Sraffa"

The Normal Degree of Capacity Utilization: The History of a Controversial Concept

Attilio Trezzini and Daria Pignalosa

Centro Sraffa Working Papers
n. 49

April 2021

ISSN: 2284-2845
Centro Sraffa working papers
[online]

The Normal Degree of Capacity Utilization: The History of a Controversial Concept

Attilio Trezzini & Daria Pignalosa

Roma Tre University

Abstract

The determination of the normal degree of capacity utilization has been addressed in the analysis of radically different problems and with different theoretical approaches. By going back over the history of these analyses, the paper highlights the main findings that have emerged from the early literature on the subject and identifies some generally shared principles on the one hand and different methodological positions on the other. Finally, attention is drawn to some issues that seem susceptible to further investigation.

Keywords: capacity utilization; choice of technique; demand-led growth; competition.

JEL codes: B51; D21.

1. Introduction[1]

The definition of a technique of production necessarily implies the specification of the intensity of use of the means of production that constitute the plant. The notion of normal utilization of productive capacity is therefore relevant in the theoretical fields in which techniques of production are crucial: the theory of prices, where normal utilization characterizes the methods of production implicit either in the production function or in the price equations, and the theory of growth, where it determines the desired capital-output ratio and the decisions to invest. In particular, in the ongoing debate within the theories of demand-led growth, the notion of normal utilization plays a critical role, so that in recent years there has been a growing stream of analyses aimed at determining this variable. These contributions largely overlook previous works that have brought to light the essential elements for determining the normal degree of utilization. Consequently, it seems worthwhile to provide a reconstruction of the early literature on the subject.

[1] We wish to thank Roberto Ciccone, Heinz D. Kurz, and Paolo Trabucchi for very helpful discussions. Any remaining errors are our own.

The variability of the degree of utilization of productive capacity has attracted the interest of scholars since the early spread of the factory system, which aroused the concern for the question of the length of the working day. However, within this vast literature, we will leave aside the many contributions that deal with the issue because of its social impact as well as those that address it because of its consequences on economic growth, focusing instead on the analyses that deal — although sometimes incidentally or even implicitly — with the *determination* of the degree of utilization planned by the firm. Such analyses are clearly widely heterogeneous, both in terms of their object and from the point of view of the theoretical approach adopted. As a first step, therefore, it is necessary to clear the ground of a series of possible ambiguities, by providing some definitions against which the different ones found in the literature can be read (section 2). We will then see that the general terms in which the problem of determining the degree of utilization is usually set, together with the basic principles for the treatment of the issue, can be traced back to Marx's analysis (section 3). However, we identify the first formal, and yet implicit, determination of the normal degree of capacity utilization as taking place within the debate on cost curves of the 1920s and 1930s (section 4). While in historical reviews these early treatments are usually neglected, we find them to raise the fundamental issues around which we will organize our entire reconstruction (section 5). A subsequent strand of analyses, originating in the 1960s, is aimed at extending the neoclassical theory of the firm (section 6). In this framework, Marris (1964) provides a detailed analysis of the role of the different elements of cost that affect the firm's choice of the planned degree of utilization. A different perspective is adopted in the early works addressing the issue within the analysis of demand-led accumulation (section 7). The contribution of Steindl (1952) inspired a later approach to the issue (Ciccone, 1986) that highlights the role of expected demand fluctuations in the determination of the size and planned degree of utilization of newly installed capacity. Within the same theoretical framework, Kurz (1986) emphasizes the need to address the issue in the context of the choice of technique.

The multiplicity of analyses considered, while not allowing to deal with each of them in an exhaustive manner, enables us to identify the main factors determining the firm's choice as to the size and the degree of utilization of productive capacity (section 8). We will also highlight, on the one hand, a generally accepted principle according to which the determination of normal utilization must be associated with profit maximization, and, on the other hand, a divergence of views on how to deal with the variability of the output produced by the firm. Furthermore, in connection with these two points, we will emphasize the link, not sufficiently investigated in the literature, between the determination of normal utilization and the definition of the market structure in which the firm is supposed to operate. Indeed, both the relevance of the variability of output and the definition of the problem of profit maximization critically depend on the notion of competition adopted, which in the contributions considered here is that of *perfect* or, alternatively, *imperfect* competition proper to neoclassical theory, or, finally, that of *free* competition of classical economists.

2. Some preliminary definitions

A source of ambiguity that appears to stem from the multiplicity of fields of analysis in which the notion of normal utilization has been studied is the lack of a generally shared definition of the variables under consideration. It seems appropriate, however, to remove such an element of heterogeneity. To this end, we will identify productive capacity with the technically maximum output obtainable through a set of durable means of production and the non-durable means of production and labour necessary to run them at the highest intensity. Accordingly, the degree to which the firm *plans* to use newly installed capacity, i.e. the normal degree of utilization, will be understood as the ratio between the level of output that the firm expects to produce on average over the life of the plant and the maximum level of output defining productive capacity. The normal — sometimes called "desired" or "planned" — degree of utilization is thus an *ex ante* measure and it implies a deliberate choice to keep spare capacity and hence an under-utilization (compared to the maximum achievable) as an *expected* result. On the contrary, actual utilization is an *ex post* magnitude, determined by the levels of output actually produced, which do not necessarily coincide with those expected. The actual degree of utilization is clearly lower or at most equal to full utilization and may be higher, lower or equal to normal utilization.

The firm can achieve a higher degree of utilization of productive capacity by increasing either the *duration* or the *speed* of operation of the plant. An increase in the duration consists in raising the number of hours per day (or week) in which the plant is run, with a given hourly product. This could involve either extending the length of the working day (resorting to overtime labour) or adopting a multiple-shift mode of operation. An increase in the speed results in a higher intensity of operation of the plant, and thus in a larger hourly output for a given number of hours of activity. This occurs when the number of workers associated with each plant increases and when each worker operates more intensely.

The maximum output, which defines the size of the productive capacity, is that produced when the plant is run 24 hours a day at the maximum intensity.

3. Marx's insights on the determination of capacity utilization

Marx never directly addresses the issue of the degree of utilization of productive capacity, yet he stands out as a pivotal reference for the literature on the subject[2] because, in various passages of *Capital* (1867-1894), approaching different issues, he draws out a number of insights that lay the foundations for modern analysis.

As a first point, Marx highlights that the ratio between the level of output produced and the fixed capital used in production is very elastic and changes following changes in

[2] See, e.g., Kurz (1986), Ciccone (1987 and 2011), and Nikiforos (2013).

"market conditions" (vol. II, p. 357), thereby recognizing the role of the variability of the degree of utilization as a means to cope with fluctuations in demand.

According to Marx, utilization can be increased "by prolongation or intensification of the working day" (vol. III, p. 326). In other words, Marx clearly identifies the two sources of variability of the degree of utilization to be found in modern literature, the first linked to the duration of operations and the second linked to the speed. Increases in duration are achieved by "exceeding the limits of the normal working day" (vol. I, p. 368), i.e. by resorting to overtime labour, or by adopting a multiple-shift system: "since it is physically impossible to exploit the same individual labour-power constantly, during the night as well as the day, […] an alternation becomes necessary, between the labour-powers used up by day and those used up by night" (vol. I, p. 367). As for increases in speed, these are accomplished when "the speed of the machines is increased" and when "the same worker receives a greater quantity of machinery to supervise or operate" (vol. I, p. 536).

A fundamental contribution of Marx's analysis is the principle according to which the determination of the (normal) degree of utilization of productive capacity must be addressed in terms of profit maximization. Furthermore, Marx singles out all the main cost components to be taken into account. First of all, he points out that the hours of overtime labour can be better paid than the regular hours, thereby identifying a very important element determining normal utilization, i.e., the wage differential: "overtime […] is paid at a better hourly rate ('extra pay'), although often in a proportion which is ridiculously small" (vol. I, p. 687; cf. also vol. III, p. 170). According to Marx, both the wage differential and the length of the regular working day depend on the same institutional and social circumstances that determine the normal wage rate. Moreover, he singles out another element of cost to be taken into account by realizing that the quantity of circulating capital used in production changes with the degree of utilization (vol. II, p. 335). Such variation of "other prime costs" — as subsequent literature will call them — is accompanied with changes in labour costs when duration increases, while it can be the only source of cost variation when increases in output are due to the intensification of speed. Depreciation is a further cost element that Marx understands to require consideration, identifying two effects of opposite sign: on the one hand "physical deterioration" (vol. I, p. 528) is likely to increase with utilization, thereby reducing the incentive to utilize capacity intensively; on the other hand, precisely because higher utilization implies a more frequent replacement of fixed capital, i.e. capital is "reproduced in a shorter series of turnover periods" (vol. III, p. 170), it lowers the risk of "moral depreciation" (vol. I, p. 528), that is technological obsolescence.

We can therefore conclude that, although his contribution does not include what can properly be qualified as a determination of the normal degree of capacity utilization, Marx provides all the elements for setting up such an analysis: the need for the determination of normal utilization to be conceived as a problem of choice of the technique that maximizes profit; the different ways in which utilization can be modified; the main cost components that are affected by a change in utilization.

4. Normal utilization in the debate on cost curves

As is well known, neoclassical analyses in which capital is treated as a homogeneous productive factor measured in value are fatally affected by problems of logical consistency[3]. However, not only by neglecting them would our treatment show a major shortcoming, it would also fail to highlight significant insights and some key controversial issues emerging from such literature.

The notion of capital as a single "quantity" inevitably hampers a clear account of the specific physical dimension of capital goods. In particular, when the assumption of variable coefficients of production is associated with the identification of productive methods by the ratio between the quantity of capital and the quantity of labour, the switch from one technique to another does not tell us by itself whether what is occurring is a variation in the intensity with which a specific plant is operated — a change in the degree of utilization of a given productive capacity — or a modification of the whole set of capital goods employed — a change in productive capacity itself. We know, however, that in long-run analyses the degree of utilization is by definition normal and thus any change in the proportion of capital to labour corresponds to a change in the physical composition of capital goods, with each good being used at its optimal level. In short-run analyses, instead, the assumption of a given productive plant implies that the only conceivable variations of output are those obtained through different degrees of utilization. That is why the first neoclassical contributions in which, although indirectly, the issue of capacity utilization is dealt with are those nourishing the debate of the 1920s and 1930s that led to the modern formulation of the short-run curves of average and marginal cost of the firm.

The "cost controversy" developed out of the dissatisfaction with the foundations of the supply curve provided by Marshall, with the now canonical cost curves emerging as a result of a "painstaking process of discussion and refinement" (Keppler and Lallement, 2006, p. 734) aimed at overcoming the ambiguities and compromises present in Marshall's analysis. While it is unnecessary to enter into the debate here,[4] it seems important to highlight the elements related to the object of our reconstruction that emerge from it. Indeed, though the analyses of cost curves are generally neglected in the literature on capacity utilization, they clearly deal with the issue as they rely on the assumption according to which any particular plant is capable of producing different output levels. Moreover, as they aim to determine how intensively the given plant must be used, such analyses appear to represent a first formal — albeit implicit — treatment of the determination of the degree of utilization chosen by the firm.

We can refer to the systematization presented by Viner (1931), who provides the now conventional representation of the curves of average fixed, variable ("direct"), and total

[3] We are referring to the results of the "capital controversy" originated from the contribution of Sraffa (1960). For a detailed account of the debate see Lazzarini (2011) and Fratini (2019).

[4] See Aslanbeigui and Naples (1997) and Marchionatti (2001). Keppler and Lallement (2006) provide an accurate account of the history of the U-shaped average cost curve, tracing the antecedents to the now classical representation provided by Viner (1931) in contributions by Barone, Edgeworth, Sraffa, Schultz, and Pigou.

costs for a firm operating in perfect competition in the short run (respectively AFC, ADC, and ATUC in Figure 1). Viner identifies the size ("scale") of the plant by "the amount of output which can be produced at the lowest average cost possible at that scale" (p. 26) — output M in Figure 1. However, according to the definition of capacity adopted here, the size of the plant is identified by the output level at which the (total or variable) average cost curve becomes a vertical line.

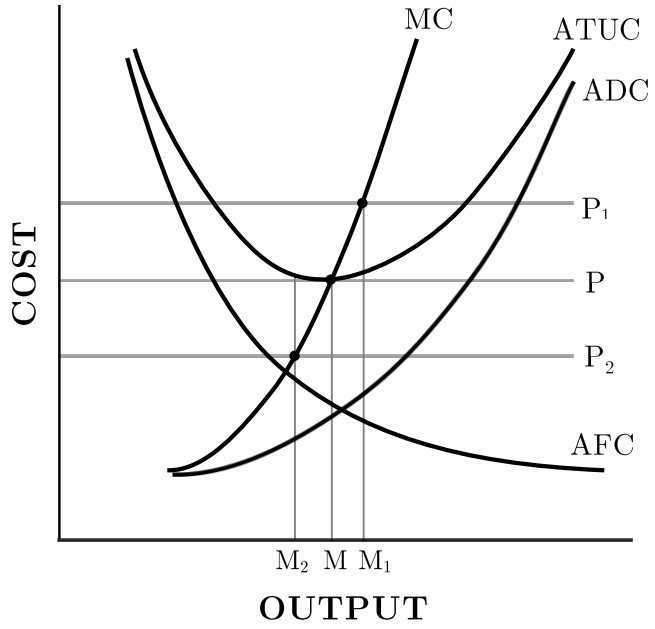

Figure 1 − Short-run cost curves (chart 1 in Viner, 1931).

As is well known, the firm will maximize profits through the equalization of marginal cost and price so that utilization will be higher than the optimal (cost-minimizing) one when the price exceeds the lowest unit cost (as is the case for price P_1) and vice versa (as for price P_2). As we shall see in the next section, such distinction between cost-minimizing and profit-maximizing utilization will prove very important for our reconstruction. However, short-term analysis is not in itself relevant for the determination of normal utilization since the latter, implying the concurrent choice of the size of productive capacity, necessarily refers to long-term conditions. Under the assumption of perfect competition, in the long run the price coincides with the minimum level of average costs, so that normal utilization is both profit-maximizing and cost-minimizing.[5]

It is important to stress that this analysis only takes into account variations in the degree of utilization resulting from changes in the speed, neglecting the possibility of varying

[5] Viner (1931), with a now legendary mistake, devised the long-run average cost curve as passing through the minimum points of the short-run functions and at the same time never lying above them. Almost simultaneously, the long-run average cost curve was more accurately derived by Harrod (1931), who properly thought it as the envelope of the family of short-run functions. However, it was only later that the relationship between short-run and long-run functions was formally described, through the "Wong-Viner" envelope theorem, by Samuelson (1947). A concise reconstruction is provided by Maclachlan (2010); for a discussion of the theoretical and empirical failures of such analysis see Brondino and Lazzarini (2017).

the duration. The shape of the cost curves outlined by Viner is indeed based on the assumption that different quantities of variable factors are associated with the fixed factor *for a given time of operation* of the latter. If the change in the quantity of labour associated with the given productive plant were instead achieved through variations in the length of the working day, then the capital-labour ratio would remain constant at the level corresponding to the maximum average product until the highest possible number of hours of activity of the plant were reached. In other words, variations in the duration of operations would manifest themselves in a horizontal section of the marginal cost curve.[6] The literature on capacity utilization of subsequent decades focuses primarily on the role of duration and it is perhaps for this very reason that it fails to consider the debate on the shape of cost curves as its own historical precedent.[7]

Another feature that distinguishes the debate on short-run costs from the following literature is that optimal utilization is determined without making explicit its connection with the choice concerning the size of the plant. In this respect, a notable exception can be identified in Stigler's (1939) contribution, which seems particularly relevant to our reconstruction. The reason why Stigler is concerned with the choice of the plant is that he takes into explicit account the hypothesis of "adaptability". Such hypothesis implies that the plant lends itself to be operated in combination with different quantities of the variable factors and thus constitutes the foundation for the standard behaviour of the marginal cost. Stigler points out that different plants are characterized by different degrees of adaptability and, moreover, that they can be more or less divisible, so that the unit cost associated with output levels other than optimal depends on the specific combination of adaptability and divisibility of the plant and the canonical representation with rising marginal cost and U-shaped average cost (Figure 1) is only one of the possible cases.

What is most interesting for the discussion at hand, however, is Stigler's idea according to which the firm, when installing its productive capacity, takes into account the fact that during the life of the plant it will have to produce different quantities of output: "the optimum output may be *an optimum through time*, i.e., the optimum may be based on anticipated increases or decreases in the output of the firm" (p. 308, emphases added).[8]

---

[6] See Maxwell (1965) and Miller (2000). It is worth pointing out that while the hypothesis of variable coefficients of production implies the possibility of varying both the duration and the speed, in the presence of fixed coefficients of production, as the capital-labour ratio per unit of time is given, the speed cannot change. In this context, an increase in daily output can only be achieved through overtime labour or by adopting a second or third shift, so that capital and labour are employed in fixed proportions as output changes.

[7] It is remarkable that a debate dealing with the implications of Marshall's contribution should overlook an issue he had clearly pointed out. See, e.g., Marshall (1920, p. 311): "The immediate effect of the expectation of a high price is to cause people to bring into active work all their appliances of production, and to work them full time and perhaps overtime". An exception to the widespread neglect of the question of the variability in the length of the working day within the literature of the 1920s is represented by Kahn's 1929 Fellowship dissertation, which however was only published in the 1980s, first in Italian and later in English (Kahn 1989).

[8] The subsequent literature on capacity utilization seems to have neglected what we regard as a particularly important aspect of Stigler's approach. Hastily mentioned in the review by Winston (1974), he is completely overlooked by Betancourt (2008).

Although Stigler does not state it clearly, the variability of the quantity produced seems to be due to the fluctuations occurring in the market in which the firm operates, which imply changes in the price of its output; these, in turn, result in changes in the marginal revenue that the firm must compare to the marginal cost in order to determine the quantity to be produced. This argument leads Stigler to introduce the notion of "flexibility", whereby a plant is more flexible the flatter is the average cost curve associated with it, i.e. the less rapidly unit costs increase in moving away from the optimal output level. When installing a plant, the firm can choose to make it more or less flexible: "[a]daptability can also be built into a plant, and entrepreneurs in trades where fluctuations are frequent and great will endeavor to secure flexibility in their operations" (p. 310). Stigler does not dwell on the methods for implementing flexibility, but he mentions the possibility of increasing the divisibility of the fixed plant or reducing the employment of the fixed plant relative to that of the variable factors (p. 316). In any case, "flexibility will not be a 'free good'" (p. 311): in order to lower the average cost associated with output levels different from the optimum, the firm must give up the lowest achievable unit cost. The degree of flexibility to be attached to a new plant will therefore be subject to economic assessment.[9]

Stigler illustrates the choice between a flexible plant A and a relatively inflexible plant B through the comparison of the curves of average and marginal cost associated with them (Figure 2). If production is expected to fluctuate between W and Z, the less flexible

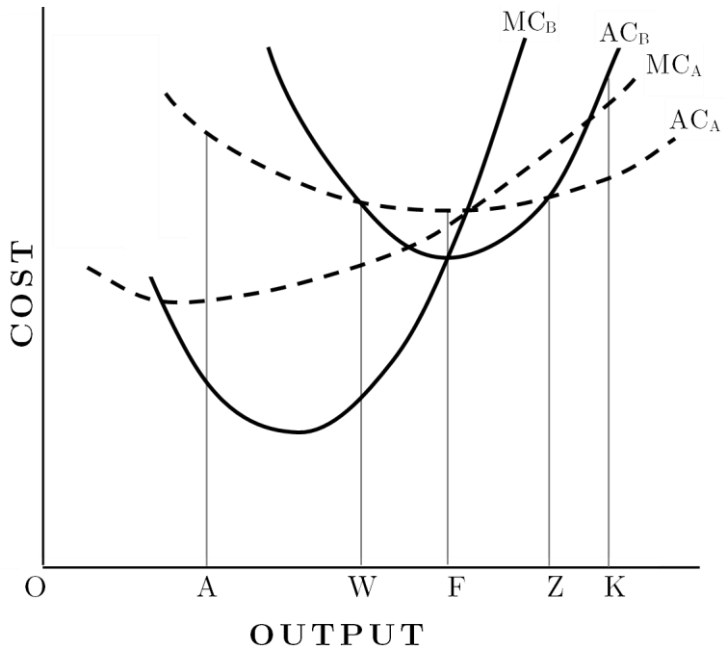

Figure 2 – Cost curves for a flexible and an inflexible plant (Chart 4 in Stigler, 1939)

[9] According to Stigler, "[e]ven after a plant has been built and equipped, it is usually possible to make alterations within short time periods which will better adapt the plant and equipment to contemporary (non-optimum) rates of output" (p. 311). Such "short-run alterations", which imply an *ex post* flexibility in addition to that implemented at the time of installation, might be important "[i]f major changes of output are frequent, relative to the life of the plant, and if they are not perfectly anticipated" (p. 319).

plant (solid lines) will be chosen, since it allows to produce all the output levels within the range at lower unit costs. When, on the other hand, the firm expects the variability of output to be sufficiently large, as is the case for quantities fluctuating in the A-K range, then the more flexible plant (dotted lines) will be desirable.

Before discussing the implications of Stigler's approach, we need to point out that, in comparing the different profits realized by the firm with a more flexible and a less flexible plant, he makes a meaningful mistake. Stigler specifies in a footnote that "[o]utputs, rather than selling prices, are selected for analysis, in order to avoid the complications arising out of the question of the slope of the demand curve if imperfect competition is considered" (p. 317). Therefore, the analysis would be applicable to both perfect and imperfect competition and the choice of dealing with fluctuations in quantity rather than price would be of a purely expositive nature. However, carrying out the comparison between the two plants taking the output levels as a starting point turns out to be misleading under the assumption of perfect competition. For, in that case, once the price has changed, the firm reacts by changing the quantity produced so that marginal costs equal the new price and, therefore, the variation in quantity following a variation in price depends upon the shape of the curve of marginal cost. In other words, comparing the profits realized with the two plants for quantities fluctuating in a certain range is pointless because a specific range of variation of price results in different ranges of variation of quantities depending on the plant that has been installed. This does not call into question the implications of Stigler's analysis on the determination of capacity utilization, but reveals that his argument would have had a sounder basis had it not been tied to the assumption of perfect competition, where it is illegitimate to suppose that firms face quantity constraints.

5. The relevance of output variability and of the notion of competition adopted

We may now make explicit the implications of Stigler's analysis with regard to the object of our reconstruction. If the choice of the plant is carried out, as Stigler suggests, taking into account that the level of output will be fluctuating, what we would call the normal degree of capacity utilization is neither identified by the quantity that the plant allows to produce at the lowest unit cost (which determines the optimal utilization) nor by any of the single quantities taken into consideration by the firm in the process of choice. Indeed, the degree of utilization that the firm plans to realize emerges as *the average of the different quantities* that are expected during the life of the plant.

The peculiarity of Stigler's argument with respect to the conventional presentation of cost curves suggests a key perspective from which to interpret all subsequent literature: since the determination of capacity utilization conducted taking into account output fluctuations (*à la* Stigler) is fundamentally different from the one in which it is assumed that the firm will constantly produce the same quantity throughout the life of the plant (*à la* Viner), the contributions aimed at identifying the normal degree of utilization may in general be classified according to the hypothesis on the variability of output that is

adopted. A firm that installs a plant intended to produce a single quantity, productive cycle after productive cycle, will ensure that such quantity is produced at the lowest possible unit cost, with no concern for the cost associated with output levels other than the optimal one. By contrast, if the firm takes into account multiple output levels, these will affect the choice of the plant, which will be sufficiently "flexible", to use Stigler's terminology, or, more generally, which will be able to produce efficiently enough output levels differing from the average quantity that identifies the normal degree of utilization. The choice of the size of the plant carried out in this way may not be profit-maximizing for any of the individual quantities of the interval (nor even for the average), but will be so with respect to the entire set, so that the normal degree of utilization, while being expressed by a single value, will depend not only on the mean of the quantities to be produced during the life of the plant, but also on the breadth of the fluctuations, provided that a wider range of oscillation of output will tend to result in a plant of greater size.[10]

That the determination of the normal degree of utilization of productive capacity necessarily implies an assumption on the variability of the output produced during the lifetime of the plant is also important because it brings to light another major issue that should be given consideration when approaching the literature on the subject. Indeed, the relevance of output variability does not seem to be independent of the notion of competition adopted. Viner's analysis is directed at the firm operating in perfect competition, which can produce whatever quantity at the given price. This notion of competition translates into a total irrelevance of demand and, as a consequence, into the assumption that the firm expects to produce a single quantity — and to continuously achieve normal utilization — over the lifetime of the plant. This case is actually peculiar, because not only is the quantity to be produced constant, it is also determined endogenously: it is through the choice of the optimal (cost-minimizing) plant that the quantity that the latter will produce is determined. As regards Stigler, he states at the beginning of the analysis that "[t]he discussion is directed primarily to perfect competition, although it is also appropriate, with suitable qualifications, to imperfect competition" (p. 306 fn), thereby suggesting that the need to produce variable output levels would arise regardless of the market structure. However, under perfect competition, the firm is affected by the fluctuations occurring in the market in which it operates only to the extent that these involve changes in the price of the product it sells. What Stigler seems to be attempting, in fact, is forcing into the theory of the competitive firm an element of reality — the influence of demand on the choice of the firm — that clashes with the neoclassical notion of perfect competition. A remarkable clue pointing in this direction can be detected in the error he makes in representing the comparison between a more flexible and a less flexible plant. That error arises because the hypothesis by which the firm chooses the plant on the basis of its *expectations* on the quantity to be produced (be it one or several quantities) is inconsistent with the assumption of "perfect" competition. The relevance of output variability can more conveniently

[10] Although Stigler's diagram does not indicate the maximum output level that can be produced by each plant, it is apparent that the average cost curve of plant B becomes asymptotic for a lower quantity than that associated with the more flexible plant.

be taken into account once, having abandoned the neoclassical notion of perfect competition, it is acknowledged that the firm has to accommodate fluctuations in the demand addressed to its product.

6. Marris' *Economics of Capital Utilisation*

In the 1960s and 1970s, the study of the determinants of capacity utilization is very extensive. The spring of interest in the issue stems from the empirical work by Foss (1963), which shows that U.S. manufacturing is characterized by persistently low levels of capital utilization, and from the theoretical contribution of Marris (1964), which stands out as the first attempt at providing an analytical explanation for *planned* capital idleness.

As he aims to explain why firms intentionally leave capital idle most of the time, Marris does not measure the degree of utilization in terms of the level of output produced, but through the number of operating hours of the equipment, with the maximum being continuous operation through all 24 hours every day. Moreover, he conceives the variability of capital utilization as stemming exclusively from the possibility of changes in the duration of operations, i.e. from the possibility of adopting alternative plans of overtime or shift-working, each corresponding to a specific *system of operation*.

The analysis is initially addressed to the case in which a single technique of production is available and the given level of output to be produced requires only one "activity". As we shall see, in this simplified scenario the total output produced is not actually a given on the basis of which the firm chooses the size of capacity and its utilization, for it is endogenously determined as the quantity that, with the given technique, yields the maximum profit. In Marris' perspective, however, the analysis of this case only represents the first step, his intention being to take into account the role of the demand addressed to the firm. Subsequently, the analysis will therefore be extended to more general hypotheses, firstly considering alternative techniques of production, then assuming that the final product requires more than one activity, and, finally, taking account of constraints on the quantity to be produced. However, the simplest model developed by Marris includes many of the most important insights and, furthermore, it is broadly independent from the neoclassical treatment of capital.

6.1. The case of one activity and a single technique

Marris takes into account that the investment decision of the firm involves the joint determination of the size and utilization of capacity ("the amount of plant, and hence the capital outlay, depends intimately on the intended rate of utilisation", p. 26), so his starting point is the definition of a *unit of capital* as the quantum of equipment corresponding to one worker and of a *unit of equipment* as the set of units of capital each of which cannot be operated independently, but which, taken together, can carry out one *activity*. An ac-

tivity is a process "with distinguishable inputs and outputs" capable of autonomous operation. When the productive process requires only one activity and a single technique of production is available to carry it out, the whole final product is produced by one specific type of equipment and what needs to be determined is the number of units to be installed, or, which is the same, the rate of utilization of one unit. To this end, Marris defines the gross rate of profit earned in one year on a specific unit of equipment with a specific system of operation as:

$$r = (mh)/C,$$

where m is the hourly profit margin, i.e. the amount of gross profit earned per hour of operating time, h is the number of operating hours per annum (the rate of utilization), and C is the capital cost of the unit. The comparison between any two alternative systems for operating the given type of equipment is to be made on the basis of the "incentive to utilize", ṙ, which is the proportionate change in the rate of profit which occurs when moving from the system of lower to the system of higher utilization. Clearly, a positive value of ṙ indicates that the system of higher utilization is the most profitable and vice versa. Since, by assumption, the firm takes input and output prices as given, C is exogenous and the change in the rate of profit when utilization is increased is $\dot{r} = \dot{h} - \dot{m} - \dot{h}\dot{m}$, which means that the incentive to utilize depends on two conflicting influences: the increase in the rate of utilization, ḣ, and the fall in the profit margin, ṁ. Therefore, "the kernel of the problem lies in those factors which determine the extent of the fall in the profit margin for a given increase in rate of utilisation" (p. 34). Now, the profit margin is given by the difference between the value of the output per unit-hour and the hourly cost of labour and circulating capital: $m = Po - n - w$. The selling price of the product, P, is given, and the hourly rate of output of the unit, o, does not vary between systems. If the hourly rate of other prime costs, n, is supposed to be constant as well, the fall in the profit margin is entirely due to the increase in the hourly wage bill w and depends on ẇ, the *utilization differential*, and w′, the ratio of wages to profits in the lower-utilization system: $\dot{m} = \dot{w}w'$. The utilization differential measures the proportionate increase in hourly wage rate caused by moving from the system of lower to the system of higher utilization: the greater the utilization differential the greater the fall in the profit margin and the smaller the incentive to utilize. The impact on the profit margin of any given utilization differential depends on the proportion of the value of output represented by wages: the higher the ratio of wages to profits the lower the incentive to utilize. In general, then, assuming that n is constant and ignoring depreciation, the general level of the incentive to utilize will depend on the general level of utilization differentials and on the factors influencing the ratio of wages to profits, i.e. on the general level of wages, on hourly productivity, on the price of the product, on the ratio of input prices to output price and on the quantity of input required to produce a unit of output.[11]

Optimal utilization thus depends on distribution: "a rise in the general level of wages, unless offset by a corresponding increase in selling price or productivity, must lower the

[11] Analytically, the ratio of wages to profits is $w' = \frac{w}{m} = 1/\left(\frac{Po}{w} - \frac{n}{w} - 1\right)$.

general level of the incentive to utilize" (p. 36). The influence on the rate of utilization of distributive variables, that is of the basic wage rate and the set of utilization differentials, is governed by the technical and cost elements that characterize the specific productive process. Among these, according to Marris, an important role can be played by other prime costs. Indeed, while initially assuming n to be constant, Marris then considers the case in which it varies between systems, pointing out that in many industries the latter decreases with utilization because of economies in fuel and raw materials that can be obtained with continuous or semi-continuous operation. In such cases the incentive to utilize is higher, to an extent that depends on the ratio of other prime costs to profits: "in an industry where other prime costs are a heavy element in total costs, economies in these costs, achieved by shift-working, will be well worth while" (p. 38).

Of particular interest is also the analysis of the role of depreciation, which Marris regards as consisting of two components. One, z, is due to factors such as rusting, rotting, and technical obsolescence, which do not vary between systems of operation. The other, u, is instead a component related to wear and tear or "user cost". When depreciation is taken into account, r can be defined as the rate of *net* profit:

$$r = h\,(m/C - u) - z,$$

where u is the proportion by which the value of the equipment depreciates, due to user cost, per hour of operating time and z is the total annual cost of other depreciation, both expressed as a ratio of C. If depreciation for wear and tear is proportional to the hours of operating time, that is, hourly user cost is constant between systems, then the higher the value of u relative to profits the lower the incentive to utilize.[12] However, according to Marris, it is likely that u decreases as utilization rises, thereby behaving like other prime costs. What is noteworthy, in the latter case, is the interaction envisaged by Marris between the two components of depreciation: "the amount of value lost through wear and tear depends on the amount already lost through obsolescence" (p. 39). In fact, what in many cases happens is that wear and tear is slight during the early stages of the machine's life and, after a certain point, increases very quickly. If the surge in wear and tear occurs when the machine has already become valueless due to obsolescence, then user cost can be entirely ignored in determining optimal utilization. Hence, the more rapid is obsolescence (i.e. the faster is technical change), the less significant is user cost and the higher the incentive to utilize.

It seems now appropriate to point out that the determinants of utilization identified by Marris clearly recall Marx's analysis. In fact, Marris seems to provide an analytical treatment to the insights that we discussed in section 3: we can find the setting of the problem in terms of profit maximization, the influence of distribution and in particular of the wage differential, the role of other prime costs and the contrasting effects of the different components of depreciation.[13]

[12] When depreciation is introduced, the incentive to utilize is $\dot{r} = \dot{h}\,[1 - \dot{m}/(1 - u')] - \dot{m}/(1 - u')$, where $u' = uC/m$.

[13] In an earlier contribution, Alexander and Spraos (1956) also highlight the same factors, but they do not provide an equally detailed analysis and, more importantly, they adopt a short-run perspective.

6.2. The case of one activity and alternative techniques

Having determined the optimal degree of utilization of a unit of equipment as the outcome of the interaction of the various factors described so far, Marris turns to the case in which there are different techniques of production to carry out the given activity. This implies that, when installing the productive capacity, it is necessary to choose not only the number but also the *type* of units of equipment. The problem is now that of determining the optimal "activity programme", which includes the technique of production (the type of equipment) and the degree of utilization (the system of operation) to be adopted. The options available to the firm can be represented through a matrix with as many rows as the number of alternative techniques and as many columns as the number of possible systems of operation. The latter are ordered from left to right from the lowest-utilization to the highest-utilization system, while the techniques are ordered according to capital intensity, so that down along a column techniques are more and more "mechanized", namely they involve "higher capital per man". Given the basic wage rate and the various utilization differentials, for each cell of the matrix — i.e. for each activity programme — the corresponding rate of profit can be computed. What has to be determined, then, is the cell that is associated with the highest profit.

According to Marris, the choice of the optimal programme depends on distribution along the lines indicated by the neoclassical theory: as wages increase, it is profitable to switch to a more capital-intensive programme, either by moving downwards, through the adoption of a more mechanized technique, or by moving left, adopting a lower-utilization system. For a sufficiently low wage rate, the programme with maximum utilization and minimum mechanization (top right of the matrix) is the most profitable, while a sufficiently high wage rate will yield an optimal programme involving the most mechanized technique and the least intensive system of operation (bottom left). In general, at each level of the wage rate, there will be a combination of system and technique that will prove optimal and for any given increase in wages, it will be more convenient to adopt one of the two alternative methods to substitute capital cost for labour cost. We do not go into the details of this rather complicated part of the analysis, which relies entirely on the illegitimate hypothesis, inherent in the neoclassical conception of the productive process, according to which techniques can be ordered according to their capital intensity independently of distribution. It is worth pointing out, though, that the discussion of the choice of technique leads Marris to suggest a further element that affects optimal utilization: "if an industry is faced with a production function in which capital is generally a 'good' substitute for labour, in the sense that relatively moderate increases in capital can yield substantial increases in labour productivity, the average optimum rate of utilisation of the industry, other things being equal, is likely to be relatively high" (p. 49).

6.3. The case of many activities

In the next stage of the analysis, Marris addresses the case of a factory whose productive process consists of different activities. In this case, optimal utilization is not referred to a

single unit of equipment, but to a "plant", i.e. the collection of equipments required for the production of the final product. The analysis gets much more complex because "the optimum rate of utilisation on any one type of equipment is inevitably affected by the operating conditions of all other types" (p. 28). While not altering the main determinants of planned utilization, this analytical effort allows Marris to define the optimal "plant programme", which states the technique of production and system of utilization to be adopted in every activity and specifies a vector of "unit requirements", i.e. the type and number of units of equipment to be installed for each activity. The entire plant will produce a quantity of final output equal to the least common multiple of the optimal quantities of the single activities (or some multiple of this least common multiple). Therefore, the level of output produced by the firm is the one that results from the combination of technique of production and system of operation that ensures the maximum profit.

6.4. Marris' treatment of output restraints

As noted at the end of the previous section, Marris analyses the factors that influence planned utilization by endogenously determining the quantity produced, as if the firm were free to sell any output level it wishes to produce. This means that, in determining "the plant design for a given output", he ignores "the possibility that the level of this given output might itself affect the solution" (p. 80). However, Marris acknowledges the need to remove such assumption: "[i]n real life no plant is designed for an unrestrained output […] A plant design must […] take account of prospective average demands and supplies over the plant life" (p. 80). As a matter of fact, taking account of output restraints is essential according to Marris, because it is mostly to them that the low levels of utilization observed should be ascribed.[14]

In the unlikely event that the exogenously given quantity that needs to be produced is a multiple of the optimal quantity associated with the plant, the firm can adopt the optimal programme; but, if the unit requirements that would allow to realize the given output are not all whole numbers but, at least partially, fractions, then the optimal programme will be unfeasible. Clearly, the problem that arises is that of the *indivisibilities* of the equipment, which are irrelevant when output is unrestrained, but, when the latter is given exogenously, force to give up the maximum profit. The firm will have to make some adjustments to the "theoretical" optimal programme in order to obtain a "practicable programme", either choosing a different technique of production compared with the one that would be optimal if there were no indivisibilities, or underutilizing the equipment (or both). In the first case, claims Marris, the technique must be "degraded", so that each unit requirement is raised to the next higher whole number. In practice, the firm can, for example, reduce the quantity of some of the capital goods which form the plant or "simply

[14] "We shall discover that restraints on output, due to market limitations, appear to be the dominant factors [in the planning decision] in the United Kingdom" (p. 122).

allow both men and machines to operate more slowly" (p. 82).[15] The alternative solution is to stick to the optimal technique but associate it with a lower-utilization system of operation, which is usually better since reductions of utilization, while raising capital costs, have the benefit of lowering labour costs.

This part of the analysis appears somewhat unconvincing, in particular because of the role of the alleged "degradation" of techniques. However, what Marris concludes is that "the general tendency of effect of restraints on output must inevitably be in the direction of reducing optimum practicable rates of utilisation" (p. 83).

Marris then wonders whether the firm facing output restraints, instead of renouncing the optimal programme, may produce a quantity that is lower than the output level exogenously given and able to yield the maximum profit. The question is in other terms whether the constraint implies a quantity that the firm must necessarily produce or it should rather be understood as a ceiling to the expansion of sales. The digression is significant, for it suggests once again that the determination of the normal degree of capacity utilization requires an explicit identification of the assumptions that establish objectives and constraints of the firm and, ultimately, define the structure of the market in which the firm operates. On this point Marris does not take any clear position: all one can find is the hypothesis of "sufficient competition" (p. 117) for firms to be price takers, and some references to unspecified "imperfections of competition" (e.g. p. 91 and p. 101) and to the firms' willingness to "strengthen their general competitive position" and "increase their share of the market" (p. 58).

6.5. Marris' approach to output fluctuations

As loosely defined as it may be, the competitive environment envisaged by Marris is certainly such as to make demand oscillations relevant for the firm's investment decision: "the plant must be so designed to be able to produce at an annual rate sufficient to meet the peak rate of demand whenever necessary" (p. 95). After all, that Marris considers the variability of output as a factor that influences the determination of planned utilization is apparent from his definition of the latter as "the middle, in some sense or other, of the *range of rates* within which [the plant] is expected to operate" (p. 7, emphasis added). It is the occurrence, during the life of the plant, of different levels of production that provides the basis for the conception of planned utilization as an average.

When Marris finally gets to the problem of taking into account demand fluctuations, he initially assumes that only two levels of demand are experienced by the firm: a certain proportion of the total output of the plant is produced at the "peak" rate and the remainder at a "normal" rate. What has to be established, then, is whether the peak should be met by adding an extra shift or by activating machinery normally kept idle. According to Marris, this amounts to determine "what rate of utilisation planned for the peak rate of output

[15] Marris is here referring to what we would call a reduction in speed; however, he does not consider it as a change in the degree of utilization but as a change of technique, for, according to his definition, variations in utilization only derive from changes in the number of hours of activity.

will maximise the average rate of profit earned by the plant over its whole life" (p. 95) and, if the proportion of output produced at the peak rate is small, then the firm is more likely to decide in favour of high rates of utilization.

It must be remarked that the perspective adopted by Marris leads him to develop his argument in largely misleading terms. Clearly, when only changes in the number of hours of operation are considered as changes in the degree of utilization, then a reserve of capacity that is expected to be left inactive most of the time does not imply a decrease in planned utilization. But, in actual fact, both if "normal" output is produced using all the machinery with a less intensive system of operation and if it is produced leaving some of the machinery idle, this amounts to a lower degree of utilization than that corresponding to peak output.

However, being aware that "in practice, the 'normal' rate of output will itself fluctuate" (p. 95), Marris then considers the case of a firm that expects a *range* of possible levels of demand and assigns a "probability of occurrence" to each of them. In such scenario, the investment plan will depend on the characteristics of the probability-distribution of demand, namely its average level, variability, and symmetry. In particular, the greater the expected degree of variation — i.e. the more unstable demand is expected to be and the higher the probabilities attached to extreme output levels — the more likely will the investment plan envisage meeting fluctuations by shifting to a higher-utilization system, rather than by maintaining a reserve of capacity. Having stated this rather generic principle, which in fact is not supported by any argumentation, Marris does not develop the analysis further. Hence, although the need to cope with demand fluctuations is presented as a key driver of firms' behaviour, the treatment of the issue appears inadequate. As a matter of fact, to the consideration of output variability Marris only devotes a hasty section, so that his contribution focuses, "almost throughout, on the question of whether, when planning to meet a given demand, one should build a plant designed to produce the required output at a low or moderate rate of utilisation, or build a smaller plant designed to produce the same output at a higher rate of utilisation" (p. 4), that is, it focuses on the case of a *single* level of output.

6.6. Subsequent developments in the neocalssical theory of the firm

With all the limitations pointed out, Marris' contribution is the first to introduce a formal treatment of the problem of determining the planned degree of utilization and to thoroughly examine the role of the various factors that influence the choice of the firm and, together with the empirical study by Foss (1963), it generated a lively research activity during the following years. For the purposes of our reconstruction, it will be sufficient to indicate the main developments and the gaps that appear to have remained unfilled.[16]

While Marris' analysis is carried out in terms of discrete techniques of production, subsequent contributions rely on a production function. By explicitly distinguishing between capital stocks and capital flows (Georgescu-Roegen, 1970), output is defined as a

[16] The early literature is reviewed by Winston (1974). See Betancourt (2008) for a more recent survey.

function of labour services (man hours per time period) and capital services (machine hours per time period), with the latter depending on the stock of capital and the rate of utilization of that stock. Furthermore, the discrete systems of utilization adopted by Marris are to be replaced by continuous-time models (Winston and McCoy, 1974) and a formal treatment of uncertainty is introduced (Smith, 1969). However, these differences do not substantially alter the main results. To sum up, according to neoclassical analyses, high capital utilization is encouraged when technology is such as to allow more capital intensive techniques and discouraged by large wage differentials, high degrees of returns to scale, and uncertainty.

As a result of the influence of Marris' work, the literature has mainly focused on explaining why firms intentionally keep plants idle for much of the time, by incorporating the choice of duration into the neoclassical theory of the firm. However, it has been pointed out that the models that adopt speed as the utilization variable produce different results because of the different behaviour of costs when utilization changes. In particular, the price of capital services increases with the speed of operations and decreases with their duration (Madan, 1987).

As for the question of output variability, this has been largely neglected. Betancourt and Clague (1981), in their extensive study of capital utilization, substantially adapt Marris' brief treatment to their set-up. While confirming that "one consequence of fluctuations in output will be to make shift-work more likely", they specify that this statement "is correct only if the elasticity of substitution is less than unity" (p. 61). A more explicit assessment of the implications of different levels of output to be produced is carried out by Smith (1970), who shows that planned capital utilization (understood as the speed of operations) is lower when the firm faces uncertain demand requirements and that an increase in the variability of demand results in an increase in the optimal capital stock.

7. Early contributions in a demand-led accumulation perspective

7.1 Steindl

Steindl's (1952) contribution "resolutely puts the concept of excess capacity and degree of utilisation in the centre of the analysis" (p. 11) and brings out the relevance of the question of capacity utilization for the theory of demand-led accumulation, in which normal utilization determines the desired capital-output ratio and hence drives investment decisions.

While the relevance of Steindl's analysis for the determination of normal utilization is largely independent of the theoretical framework adopted, it must be noted that his work is related to the debate on the theory of imperfect competition, as he is prompted by the lack of a theoretical explanation for the existence of excess capacity in oligopolistic industries. According to empirical analyses, Steindl points out, excess capacity is "just as ordinary in oligopolistic industries as in others" and, although the data may well show

some “unintended and unplanned” excess capacity, the evidence suggests that a low degree of utilization of plants and equipment is “accepted as normal by the manufacturers in those industries” (p. 6).

An important aspect of Steindl’s contribution is the distinction between the *planned* degree of utilization and the *optimal* one, with the latter being understood as that which minimizes unit costs and which is achieved, in long-run equilibrium, in perfect competition. Such distinction first requires the definition of the “practical capacity”, the measure adopted as a benchmark against which to determine the degree of utilization and that is identified by “the output achieved with normal length of working time, with sufficient shut-downs to allow for repairs and maintenance, and without disturbance in the smooth running of the production process” (pp. 7-8). Since empirical evidence suggests that firms’ marginal costs are approximately constant up to practical capacity and then increase sharply, Steindl reckons the optimal degree of utilization to be very close to the maximum. The excess capacity reported by the data, then, points to a degree of utilization that is lower than the one minimizing unit costs. Steindl’s objective is therefore to explain why firms deliberately install overcapacity, thereby choosing a “planned” degree of utilization lower than the optimum.

The determinants of the planned degree of utilization, according to Steindl, are to be found in part in the characteristics of the demand addressed to the firm and in part in those of the productive capacity through which the firm is supposed to meet that demand.

If plants were, firstly, divisible and at constant returns and, secondly, replaced frequently enough, then capacity could be adjusted to production continuously and perfectly, expanding it whenever output needs to be expanded and vice versa; but, because of its durability, each plant will be operated, in its lifetime, both during phases of expansion of the firm’s output and during phases of contraction, and if its size is chosen taking such fluctuations into account, then the plant will be underutilized when going through phases of low production. Moreover, the indivisibility of plants may force the firm to install productive capacity that results excessive even with respect to peaks in production.

The fluctuations in the quantity produced by the firm derive mainly from those of the demand addressed to the industry: the firm must be able to meet peaks in demand in order to prevent its (existing or potential) competitors from exploiting them to take over market shares. Besides, installing excess capacity may be desirable even when the demand of the industry is expected to follow a steady or even declining trend if the firm looks to increase its market share to the expense of its competitors. These arguments lead Steindl to consider as particularly important, in explaining the discrepancy between planned and optimal utilization, the role of uncertainty and, consequently, that of the subjective expectations of entrepreneurs and, more generally, of their degree of optimism. Indeed, the failure to properly account for uncertainty is, according to Steindl, one of the major deficiencies of the theory of imperfect competition.

We can now point out that Steindl’s idea according to which firms take into account that they will have to produce different levels of output as they install their productive capacity is one that is shared by Marris, but, as we noted, the latter fails to carry out the

analysis that would be necessary to draw significant implications from that premise. However, we have found such an analysis within Stigler's contribution, in which the fluctuations in the quantity produced play as important a role as that attributed to them by Steindl. Clearly, the distance in terms of theoretical approach and research objectives precludes any proper comparison between the two scholars; yet it may be pointed out that, by explicitly addressing the firm operating in an oligopolistic industry, Steindl's analysis finds its presupposition in the firm's willingness to defend and expand its market share, so that the relevance of the variability of the quantity produced is justified by a strategic element and soundly rooted in the assumption qualifying the market structure. As we have pointed out, Stigler's analysis raises several doubts in this respect.

It should at the same time be pointed out that unlike Stigler's, Steindl's contribution is not grounded on a specific description of the firm's decision-making from which the normal degree of utilization may emerge as the result of a profit-maximization process. In fact, the idea of defending and expanding the firm's market share, while being legitimized by the explicit departure from the hypothesis of perfect competition, is not developed satisfactorily when it comes to the choice of productive capacity: there is no formal analysis from which the need to accommodate peaks in demand emerges as an analytical result and which, in addition, clarifies whether the firm should aim to meet whatever peak occurs or whether it is some process of assessment of their profitability that indicates which peaks are worth meeting. Indeed, one can get the impression that Steindl's analysis lacks the generally accepted principle according to which the normal degree of capacity utilization should be the outcome of a profit maximization process. However, Steindl explicitly identifies its objective in an explanation of the excess capacity compatible with the maximizing behaviour of firms, and firmly rejects those aspects of the theory of imperfect competition "departing from the principle of maximisation of profits" (p. 2). What is missing is rather an articulated formulation of this principle that only a more formalized analysis would have allowed. As a way to reinforce his arguments, Steindl prefers to the latter the reference to the empirical evidence on excess capacity, thus leaving unspecified the different role of the various factors that affect the choice of the firm. In any case, Steindl's "planned" degree of utilization, while being non-optimal in the sense that it does not correspond to the minimum unit cost associated with the plant, is undoubtedly optimal to the extent that, given the range of output levels expected, the size and degree of utilization of capacity are the ones that allow to maximize profits over the life of the plant.

7.2 Ciccone

In 1986, Ciccone takes up and develops the approach based on the explicit consideration of output fluctuations put forward by Steindl. Since his theoretical framework is to be found in the modern revival of the surplus approach, in discussing the concept of normal utilization, Ciccone refers to the degree of utilization implicit in Sraffa's (1960) price equations. As part of a critical discussion of Joan Robinson's theory of distribution, he aims to state that the flexibility of the degree of utilization is not limited to the short period

and that, as a consequence, actual utilization can be different from normal utilization even in the long period.[17] This leads Ciccone to discuss the circumstances upon which normal utilization depends, without requiring, however, a formal determination of the latter.

According to Ciccone's definition, normal utilization is "that expected for newly installed capacity" (p. 24) and is determined by the same circumstances that affect the size of capacity, namely, in the first place, "the fluctuations that, in a market economy, generally characterize demand, and hence, more or less closely, production" (p. 26). As entrepreneurs cannot but take output variability into account in their decisions, the size of capacity will be "commensurate with the *relatively higher levels of demand* that entrepreneurs expect to encounter, with a certain frequency, during the economic life of their plant" and normal utilization will be "that which entrepreneurs expect to realize *on average*, over long periods of time, as a result of the fluctuations in the degree of utilization" (p. 27, emphasis added). It will therefore not only be considerably lower than full utilization, but "the smaller, the larger are the breadth and frequency of the expected falls in production with respect to the peaks for which capacity is adequate" (p. 27).[18]

Like Steindl, Ciccone considers the variability of output as the key factor determining planned utilization, but, taking account of Marris' contribution, he also points to the potential profitability of excess margins of capacity deriving either from the adoption of more convenient productive techniques or from the lower unit costs of production associated with a less intensive utilization.

Steindl and Ciccone also share the recognition of the rationale for the relevance of demand fluctuations "in the need of the individual firm not to lose market shares when demand goes up — and so, ultimately, in the pressure of competition" (p. 27). In this respect, it is important to remark the different assumptions adopted by the two authors with regard to the market structure, whereby Steindl refers to oligopolistic industries and Ciccone to a more general framework of free competition. In fact, the hypothesis according to which firms should prevent the appropriation of market shares on the part of (existing) competitors seems to be applicable to any context where firms' size is not atomistic[19] and the argument can be extended to potential competitors if one assumes freedom of entry into the market. Hence, while being clearly inconsistent with perfect competition, the hypothesis does not require any specific form of "imperfection" of the market. The notion of competition of the classical economists embraced by Ciccone, inasmuch as it only implies the free movement of capitals between sectors ensuring the tendency towards a uniform rate of profit, is less restrictive of both perfect and imperfect neoclassical competition and

[17] The long period is understood as "a time-span that is sufficiently long for the gravitation of prices and quantities produced around their respective normal values to manifest itself" (p. 23). On the conception of long-period or normal positions, see Garegnani (1976).

[18] Such arguments, also put forward by Garegnani (1992), represent the foundations of what has been defined as "the flexible-utilization approach" to the analysis of demand-led growth (Trezzini and Palumbo 2016).

[19] More precisely, all cases where the size of the firm is somehow constrained (e.g. by rapidly decreasing returns to scale) must be ruled out, for what is required is the possibility for the firm to expand.

compatible with the idea that firms aim at defending their market share from existing and potential competitors.[20]

7.3 Kurz

The contribution by Ciccone (1986) that has just been discussed is part of the debate on the compatibility between the concept of "normal positions" adopted by the classical economists and Keynes's theory of effective demand. Within that same debate, Kurz (1986) provides a formal treatment of the problem of determining the degree of utilization involved in the concept of normal position. In particular, his contribution is intended to emphasize that, since its determination "forms an integral part of the choice of technique" (p. 46), normal utilization results from the minimization of unit costs and, as such, it depends on income distribution.[21]

In his "exceedingly simple example" (p. 48), Kurz addresses the choice between two alternative modes of operation by considering an entrepreneur who can produce the given annual level of output either by working M machines under a single-shift system or by employing $\frac{M}{2}$ machines and double shifts. Following Sraffa's (1960) approach to the problem of the choice of technique, Kurz compares the cost of production associated with the two systems of operation. With the higher-utilization system the cost of fixed capital is lower because the same output is produced by working half of the machinery twice as long, while the cost of labour is higher because of the wage differential that must be paid to workers during the second shift. Clearly, the double-shift system will be adopted if it implies a reduction in capital costs larger than the increase in labour costs.

A set of simplifying assumptions makes the comparison between the two modes of operation easier. In particular, production is characterized by constant returns to scale and the machines exhibit constant efficiency throughout their life, which lasts two years under the single- and one year under the double-shift system. Now, under these assumptions, if p_m is the price of the (new) machines and r is the general rate of profit, the cost of fixed capital, including normal profits, is $Mp_m\frac{(1+r)^2}{2+r}$ with the lower- and $\frac{M}{2}p_m(1+r)$ with the higher-utilization system.[22] As for labour costs, if w is the basic wage rate, α the shift differential, and L_Q the amount of labour required to produce the given annual level of output, then the wage bill is $L_Q w$ under the single-shift system and $L_Q w\left(1+\frac{\alpha}{2}\right)$ otherwise.

[20] For a discussion of the different definitions of competition within classical and neoclassical analyses, see Eatwell (1982).

[21] Kurz's purpose is to dispute Vianello's (1985) and Ciccone's (1986) contention that an over-utilization of capacity could result in a higher than normal profit rate. According to Kurz, since the normal degree of utilization is that which maximizes profits, only a fall of real wages can make the switch to a more intensive utilization profitable.

[22] Under the double-shift system, machines behave as circulating capital as they are worn out entirely during the year, while for the single-shift system it needs to be recalled that if the life of a machine is n years and its value when new is p_m, at the age of t years its value is $p_m[(1+r)^n-(1+r)^t/(1+r)^n-1]$ (Sraffa, 1960).

The comparison thus leads to conclude that the double-shift system is adopted whenever $Mp_m \frac{r+r^2}{2(2+r)}$ is larger than $L_Q \frac{\alpha}{2} w$.[23] Kurz can therefore claim on the one hand that normal utilization cannot be determined regardless of income distribution, and on the other that profit maximization is compatible with "firms [planning] intentionally to leave their capital stock idle over substantial stretches of time" (p. 49).

Admittedly, the simplifying assumptions that underlie the example just depicted conceal some of the key determinants of normal utilization mentioned in the previous sections. Being familiar with both Marx's and Marris' analyses, Kurz himself acknowledges neglecting several factors whose relevance may be crucial. He mentions in particular the possibility of non-constant returns to scale, the non-proportional effect of higher utilization on the rate of depreciation of fixed capital, the effects on the pattern of maintenance and repair activities, on the efficiency of machinery and on the productivity of labour. If it is therefore hard to deny that the example proposed by Kurz is unable to encompass many of the elements involved in the determination of normal utilization, it must however be remarked that it represents an important basis susceptible of development. On the other hand, we may include in the set of particularly restrictive assumptions the one according to which the entrepreneur bases his investment decision on the supposition of having to produce a given, single level of output throughout the life of the plant. Taking this into account and considering that Kurz's and Ciccone's stances on the issue bring out the dichotomy which we have already referred to between the analyses that attach a role to the variability of output and those that do not, it seems appropriate to dwell on Kurz's choice to consider the quantity to be produced as constant.

In explicit opposition to Ciccone (1986), Kurz contends that fluctuations in demand and output (unless so regular as to be considered "normal") should not be taken into account in defining the normal degree of capacity utilization implicit in Sraffa's price equations. The lack of regularity of fluctuations and their dependence on temporary and accidental factors would make a normal degree of utilization determined on the basis of them unstable and accidental as well and would deprive normal prices of the persistence which must be the hallmark of positions operating as centres of gravitation. In determining normal utilization, therefore, it would be legitimate to take into account only seasonal or in any case regular fluctuations. In fact, Kurz ignores also this source of output variability in his example, but he acknowledges both its empirical relevance and that it can produce far from negligible effects on the determination of utilization (p. 47).[24]

[23] Kurz also shows that the "reswitching of techniques", at the centre of the debate on the neoclassical notion of capital, can also occur when considering two methods of production that differ only in the degree of utilization of the same plant (pp. 50-51).

[24] In support of his arguments, Kurz highlights two relevant elements that the literature on normal utilization has overlooked, namely the role of obsolete capital and of inventories in mitigating the variability of capacity utilization. In response to fluctuating levels of demand, Kurz suggests, the firm can bring into work machines of an obsolete type usually kept idle as well as modify the amount of (non-perishable) products held in stock, so as to keep a relatively steady flow of production and therefore a fairly steady degree of utilization of the latest vintages of capital goods (cf. Kurz, 1986, p. 47 and Kurz and Salvadori,

In connection with this, we may note that the attention that Kurz pays to price theory seems to induce him to consider only two kinds of fluctuations, which are indeed the only ones that are relevant in the analysis of price gravitation: those of a seasonal or cyclical nature and the purely accidental fluctuations. In fact, another kind of "irregular" fluctuations is crucial in accumulation analysis, namely the expansions and contractions of output of different duration and intensity that generate the more or less growing trend of overall output, i.e. the fluctuations through which growth actually takes place over time.[25]

In his reply, Ciccone (1987) argues that the instability to which Kurz refers is inherent in *actual* utilization, which varies with fluctuations in demand, whereas it does not characterize normal utilization. If actual utilization varies irregularly, entrepreneurs cannot ignore this fact when determining the size and planned utilization of productive capacity; however, this does not mean that the normal degree of utilization is influenced by current fluctuations, but only that fluctuations experienced in the past lead entrepreneurs to form expectations of variability in production. Once these expectations result in a specific planned degree of utilization, the latter, incorporated into normal techniques, will be stable enough to ensure the necessary persistence of normal prices.

8. Concluding remarks

The concept of normal capacity utilization has assumed a crucial role in the debate within the analyses of demand-led growth, becoming the focus of a growing body of literature.[26] While the variable that is at the centre of the debate is the normal degree of utilization at the aggregate level, it would seem that for solving the controversy one cannot but start from an understanding of the determinants of normal utilization at the micro level. However, the debate has developed without focusing on the analysis of the profit-maximization process that underlies the choice of the size and planned utilization of capacity. Only recently some contributions have explicitly addressed the investigation of firm behaviour (e.g. Nikiforos, 2013).

The present paper goes over the history of the early analyses of normal utilization, the results of which seem to have been largely neglected, but that, in our opinion, can provide important insights to be incorporated in the ongoing debate.

As it is not possible to give an exhaustive reconstruction of such a wide-ranging and heterogeneous literature, we have attempted to highlight the main findings that have

1995, pp. 348-351). Clearly, the possibility of resorting to inventories and obsolete capital must be evaluated on the basis of the principle of cost minimization. Moreover, the planned average level of stocks, as well as the extent to which obsolete machines are worth employing will depend on the characteristics of demand fluctuations.

[25] Cf. Trezzini (2020).

[26] To put it very briefly, the problem is that of combining an autonomous role for aggregate demand in the growth process with an exogenous determination of distribution. In this respect, it is essential to clarify whether actual utilization should be assumed to equal normal utilization in the long run and whether normal utilization should be conceived as endogenous to demand.

emerged and to draw attention to some issues that are more controversial or in any case susceptible to further investigation.

A principle that meets with general consensus is that according to which the normal degree of utilization is that which the firm plans upon installing the productive capacity and is determined in such a way as to maximize profits, so that its determination is a component of the process of choice of technique.

As is clear ever since Marx's analysis, the variability of the degree of utilization can take two distinct forms, for the firm can either modify the speed of operations or their duration. However, as we have seen, most of the literature has been concerned with determining the duration planned by the firm, while the determinants of the speed remain largely unexplored. In fact, the two dimensions should be considered jointly as they have different and sometimes opposite effects on the behaviour of the costs of production.

With regard to the normal degree of utilization understood as the duration corresponding to the maximum profit, its determination requires assessing the effect of changes in the length of operations on the average cost of fixed and circulating capital and on the cost of labour. As utilization increases, the cost of fixed capital decreases because the same plant produces a higher level of output; at the same time, a higher utilization can increase the depreciation due to wear and tear, although this effect becomes negligible when technical obsolescence is sufficiently fast. The possible impact of a higher utilization on the cost of circulating capital should also be considered, because continuous or semi-continuous operation can imply economies in fuel and raw materials. What certainly discourages an intensive utilization of plants, instead, is the higher wage that must be paid to workers during non-standard hours. Overall, therefore, the degree of utilization corresponding to the maximum profit depends on the circumstances that determine the behaviour of costs of production as duration changes, i.e. the technical conditions of production, the price system, and the distributive variables, including, clearly, the set of wage differentials.

The analysis of the literature has shown that on a relevant point two distinct positions can be identified. In a broad and heterogeneous set of studies, the determination of normal utilization is carried out with reference to a single output level, while another group of analyses assumes that the firm takes into account that different quantities will be produced during the lifetime of the plant. Almost all the analyses of cost functions of the 1920s and 1930s adopt the assumption of a single quantity, determining the one profit-maximizing output level. The same hypothesis can be found in most of the subsequent neoclassical literature on capital utilization as well as in the analyses with a demand-led growth perspective. The other position can be traced back to the analysis of Stigler (1939), who points out that when a single quantity is taken into account the cost associated with all the other output levels is irrelevant, but, if a range of quantities is to be accounted for, the plant must be able to produce all the output levels within the range in a sufficiently efficient manner. Marris (1964) attaches great importance to output fluctuations and yet he does not give any analytical role to them. It is in fact in the contributions of Steindl (1952) and Ciccone (1986) that an explicit analysis of the relevance of output fluctuations for

determining normal utilization can be found. Under the assumption that the firm aims at meeting peaks of demand for its product, they conceive normal utilization as the average of the different degrees of utilization expected during the whole life of the plant, and as such depending not only on technical conditions and distribution, but also on the features of expected demand fluctuations.

One implication of taking into account output variability is that normal utilization is unlikely to correspond to the output level that the plant could produce at the lowest unit cost, i.e. normal utilization will be different from that which is usually called "optimum". It is however important to stress that this result, which is most likely to be obtained under the assumption of a variable output, is still quite probable when the quantity produced by the firm is assumed to be constant, as long as it is exogenously given. For the indivisibility of fixed capital can result in the installation of a plant that would be able to produce at the lowest unit cost a different quantity from that expected to be produced by the firm.[27] As we have seen, the role of indivisibilities, strongly emphasized by Steindl (1952), was also acknowledged by Marris (1964), who hinted at the "degradation of technique" as a way of dealing with output restraints that prevent the production of the quantity corresponding to the maximum profit.

The relevance of the single or multiple levels of output to be produced suggests that the determination of the normal degree of utilization of capacity is crucially affected by the assumptions about constraints and objectives of the firm, i.e. by the notion of competition adopted. In this regard, one result that we have highlighted is the incompatibility of the neoclassical notion of perfect competition with the hypothesis that the firm plans the size and degree of utilization of its capacity on the basis of the output that such capacity will have to produce. The hypothesis of perfect competition implies that the firm faces a horizontal demand and, therefore, that the quantity produced is determined endogenously while normal utilization depends exclusively on technique and cost conditions. By contrast, in most of the analyses considered, normal utilization is seen as depending also on expectations about future production. As a matter of fact, while the analyses of cost functions such as that by Viner (1931) rely on the assumption of perfect competition, the contributions that explicitly address the determination of normal utilization adopt different hypotheses: Steindl (1952) deals with the case of oligopolistic industries, Marris (1964) refers to a not precisely defined context of imperfect competition, Ciccone (1986) and Kurz (1986) adopt the notion of free competition proper to classical economists. An exception is Stigler's (1939) analysis, whose weaknesses derive precisely from the attempt to make the choice of plant depend on the quantity to be produced despite the framework of perfect competition.

Overall, the analysis carried out reveals the need to explicitly address the hypotheses on constraints and objectives assigned to firms, which are often not outlined in such a detailed fashion as to clearly delineate the framework in which a formal analysis of the

[27] Alexander and Spraos (1956) refer to "the distinction between producing *any* given […] output at its minimum average cost […] and producing that *unique* output, which, out of the whole range of potential outputs, yields the lowest average cost" (p. 609, emphasis in the original).

determination of the degree of normal capacity utilization can be undertaken. Furthermore, it is important to adequately account for the implications of output fluctuations, by embedding Steindl's and Ciccone's insights into an analytical examination of the profit maximization process. The literature discussed here seems to provide the essential elements for the development of such an analysis, which, although undoubtedly complex, appears to represent a necessary step to clarify the many issues that the concept of normal utilization raises in different contexts of theoretical analysis.

References

Alexander, K.J.W. and Spraos, J. (1956), Shift Working: An Application of the Theory of the Firm, *The Quarterly Journal of Economics*, 70(4), 603-612.

Aslanbeigui, N. and Naples, M.I. (1997), Scissors or Horizon: Neoclassical Debates about Returns to Scale, Costs, and Long-Run Supply, 1926-1942, *Southern Economic Journal*, 64(2), 517-530.

Betancourt, R.R. (2008), Capital Utilization, in: S.N. Durlauf and L.E. Blume (eds), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan.

Betancourt, R.R. and Clague, C.K. (1981), *Capital Utilization: A Theoretical and Empirical Analysis*, Cambridge University Press.

Brondino, G. and Lazzarini, A. (2017), Sraffa's 1920s Critique and its Relevance for the Assessment of Mainstream Microeconomics, *Research in the History of Economic Thought and Methodology*, 35, 131-151.

Ciccone, R. (1986), Accumulation and Capacity Utilization: Some Critical Considerations on Joan Robinson's Theory of Distribution, *Political Economy: Studies in the Surplus Approach*, 2(1), 17-36.

Ciccone, R. (1987), Accumulation, Capacity Utilization and Distribution: A Reply, *Political Economy: Studies in the Surplus Approach*, 3(1), 97-111.

Ciccone, R. (2011), Capacity utilization, Mobility of Capital and the Classical Process of Gravitation, in: R. Ciccone, C. Gehrke, and G. Mongiovi (eds), *Sraffa and Modern Economics*, Volume 2, Routledge, 76-86.

Eatwell, J. (1982), Competition, in I. Bradley and M. Howard (eds), *Classical and Marxian Political Economy. Essays in Honour of R.L. Meek*, Palgrave Macmillan, 203-228.

Foss, M.F. (1963), The Utilization of Capital Equipment: Postwar Compared with Prewar, *Survey of Current Business*, 43(6), 8-16.

Fratini, S.M. (2019), On the Second Stage of the Cambridge Capital Controversy, *Journal of Economic Surveys*, 33(4), 1073-1093.

Garegnani, P. (1976), On a Change in the Notion of Equilibrium in Recent Work on Value and Distribution, in: M. Brown, K. Sato, P. Zarembka (eds), *Essays in Modern Capital Theory*, North Holland, 25-45.

Garegnani, P. (1992), Some Notes for an Analysis of Accumulation, in: J. Halevi, D. Laibman, and E. Nell (eds.), *Beyond the Steady State. A Revival of Growth Theory*, Macmillan, 47-71.

Georgescu-Roegen, N. (1970), The Economics of Production, *American Economic Review*, 60(2), 1-9.

Harrod, R.F. (1931), The Law of Decreasing Costs, *The Economic Journal*, 41(164), 566-576.

Kahn, R.F. (1989), *The Economics of the Short Period*, Macmillan.

Keppler, J.H. and Lallement, J. (2006), The Origins of the U-Shaped Average Cost Curve: Understanding the Complexities of the Modern Theory of the Firm, *History of Political Economy*, 38(4): 733-774.

Kurz, H.D. (1986), Normal Positions and Capital Utilization, *Political Economy: Studies in the Surplus Approach*, 2(1), 37-54.

Kurz, H.D. and Salvadori, N. (1995), *Theory of Production. A Long-Period Analysis*, Cambridge University Press.

Lazzarini, A. (2011), *Revisiting the Cambridge Capital Theory Controversies: A Historical and Analytical Study*, Pavia University Press.

Maclachlan, F. (2010), Long-run and Short-run Cost Curves, in: M. Blaug and P. Lloyd (eds), *Famous Figures and Diagrams in Economics*, Edward Elgar, 73-79.

Madan, D.B. (1987), Optimal Duration and Speed in the Long Run, *Review of Economic Studies*, 54(4), 695-700.

Marchionatti, R. (2001), Sraffa and the Criticism of Marshall in the 1920s, in T. Cozzi and R. Marchionatti (eds), *Piero Sraffa's Political Economy: A Centenary Estimate*, Routledge.

Marris, R. (1964), *The Economics of Capital Utilization: A Report on Multiple Shift Work*, Cambridge University Press.

Marshall, A. (1920), Principles of Economics, 8th ed., Palgrave Macmillan, 2013.

Marx, K. (1867), *Capital: A Critique of Political Economy*, vol. I, Penguin Books, 1990.

Marx, K. (1885), *Capital: A Critique of Political Economy*, vol. II, Penguin Books, 1992.

Marx, K. (1894), *Capital: A Critique of Political Economy*, vol. III, Penguin Books, 1992.

Maxwell, W.D. (1965), Short-Run Returns to Scale and the Production of Services, *Southern Economic Journal*, 32(1), 1-14.

Miller, R.A. (2000), Ten Cheaper Spades: Production Theory and Cost Curves in the Short Run, *The Journal of Economic Education*, 31(2), 119-130.

Nikiforos, M. (2013), The (normal) rate of capacity utilization at the firm level, *Metroeconomica*, 64(3), 513-538.

Samuelson, P. (1947), *Foundations of Economic Analysis*, Harvard University Press.

Smith, K.R. (1969), The Effect of Uncertainty on Monopoly Price, Capital Stock and Utilization of Capital, *Journal of Economic Literature*, 1(1), 48-59.

Smith, K.R. (1970), Risk and the Optimal Utilization of Capital, *Review of Economic Studies*, 37(2), 253-259.

Sraffa, P. (1960), *Production of Commodities by Means of Commodities. Prelude to a Critique of Economic Theory*, Cambridge University Press.

Steindl, J. (1952), *Maturity and Stagnation in American Capitalism*, Oxford University Press.

Stigler, G. (1939), Production and Distribution in the Short Run, *Journal of Political Economy*, 47(3), 305-327.

Trezzini, A. (2020), Harrodian Instability: An Unhelpful Analytical Concept, *Review of Radical Political Economics*. 10.1177/0486613420967201

Trezzini, A. and Palumbo, A. (2016), The theory of output in the modern classical approach: main principles and controversial issues, *Review of Keynesian Economics*, 4(4), 503-522.

Vianello, F. (1985), The Pace of Accumulation, *Political Economy: Studies in the Surplus Approach*, 1(1), 69-87.

Viner, J. (1931), Cost Curves and Supply Curves, *Zeitschrift Für Nationalökonomie*, 3(1): 23-46.

Winston, G.C. (1974), The theory of capital utilization and idleness, *Journal of Economic Literatur*e, 12(4), 1301-1320.

Winston, G.C. and McCoy, T.O. (1974), Investment and the Optimal Idleness of Capital, *Review of Economics Studies*, 41(3), 419-428.

Author contact information:

Attilio Trezzini
Department of Economics
Roma Tre University
Via Silvio D'Amico, 77
00145 – Rome (Italy)
attilio.trezzini@uniroma3.it

Daria Pignalosa
Department of Economics
Roma Tre University
Via Silvio D'Amico, 77
00145 – Rome (Italy)
daria.pignalosa@uniroma3.it