DISCUSSION PAPERS IN ECONOMICS
MANCHESTER METROPOLITAN UNIVERSITY

The Industry Supply Curve: Two Different Traditions

Arrigo Opocher * and Ian Steedman**

* Dipartimento di Scienze Economiche, Università di Padova, Via del Santo 33, Padova 35123, Italy

**Department of Economics, Manchester Metropolitan University, Mabel Tylecote Building, Cavendish Street, Manchester, M15 6BG, UK

Abstract. This paper seeks to provide some new insights into the precise nature and the analytical foundations (or lack of them) of the familiar industry supply curve. To this end, we reconsider some fundamental phases of its historical evolution. Two different traditions are distinguished: one consists of the formalisations of Marshall's theory proposed by Barone and, later, by Pigou, Viner, Harrod and Robinson; the other consists of the models of Hicks and Allen, on the basis of ideas and criticism put forward by other LSE scholars, like Kaldor and Robbins, in the mid-1930s. It is argued that the second tradition did not really remedy the weak aspects of the Marshallian theory of supply.

Working Paper: 2006-02

1. Introduction

Two different kinds of economic reasoning have historically motivated the familiar industry supply curve. One consists in *representing* in the two dimensional space of output and price a more inclusive relationship involving other prices and other quantities: in this conception, a movement along the supply curve involves a change in many economic variables, such as input prices and other product prices. The other reasoning simply consists in *assuming* all prices to be constant *except* that particular commodity price and then determining output (of the individual price-taking firm and of the industry) as a function of that price. The first approach was adopted and developed by Marshall and by a series of commentators in the Marshallian tradition, like Pantaleoni, Barone, Ricci and, later, Pigou, Viner, Harrod, Robinson. The second approach was prompted by a radical criticism of the first and was developed during the 1930s by *LSE* scholars, such as Robbins, Kaldor, Hicks, and Allen and was later refined by Samuelson.

This paper presents and compares the analytical underpinnings of the main contributions to these two different traditions, in an attempt to clarify the precise nature and the historical roots of the supply curve as we find it today in reputable microeconomics textbooks. It will be argued, in particular, that the LSE "revolution" did not really remedy the weak aspects of the Marshallian tradition, and that the analytical foundations of the industry long-period supply curve are still open to question.

In § 2 we present a brief sketch of Alfred Marshall's original conception, setting his mature theory against the background of his early versions. He did not fully develop a *formal* analysis and subsequent contributions can be interpreted as formalisations of one aspect or another of Marshall's "curve". The model of Barone (1992)[1894] is discussed in § 3, while the formalisations of Pigou (1928) and Viner (1953) [1931] are the main subject of § 4. All these versions of the supply curve, including Marshall's, are based on free entry and a zero pure profit condition. This was, of course, at the heart of the leading long-run theories of

1

Marshall's time. In particular, Wicksell called "full equilibrium" the situation in which the firm operates at the lowest point of a U shaped average cost curve[1], earning zero pure profits. The Hicks-Allen-Samuelson "revolution", discussed in § 5, eventually led to a different conception of the long run, which emphasizes potential closure much more than potential entry. This will lead us quite naturally, in § 6 and § 7, to consider current textbook approaches, and to ask how effectively they cope with the problems, criticisms and weaknesses which characterised the historical evolution of the supply curve. Section 8 concludes.

2. "Full equilibrium" and Marshall's supply curve

We know that Alfred Marshall developed his diagrammatic analysis of supply and demand well before the first edition of the *Principles*. His manuscripts on value (MSS hereafter) and his *The Pure Theory of Domestic Values*[2] (PTDV hereafter), both written in the 1870s, contain a vivid exposition of the main theoretical arguments, which were to be developed, with the addition of many factual illustrations, in Book V of the *Principles*: it is therefore useful to consider such versions as a background to his mature construction.

---

[1] See Wicksell (1935)[1901], pp. 129- 130.

[2] The manuscripts on value, published with extensive historical commentary by Whitaker (1975, vol 1, pp. 119-164), were probably prepared by Marshall for his "Tripos lectures" at Cambridge in 1870 or so (Groenewegen, 1995, pp. 154-5). A more refined version belonged to the project of an early unpublished treatise on foreign trade, whose draft was nearly completed in 1875-7. The leading European economists saw Marshall's diagrams in two pairs of chapters, taken from the planned treatise and privately printed by H. Sidgwick in 1879, under the titles of *The Pure Theory of Foreign Trade* and *The Pure Theory of Domestic Values* (eventually published posthumous by the London School of Economics: see Marshall, 1930). They soon became a sort of common knowledge, a shared acquisition, to the extent that some of them first reached the general public, with acknowledgments of course, in an Italian text, Pantaleoni's *Principii di economia pura* (1889).

According to Marshall, an industry which brings to the market a certain regular "yearly" flow of output, operates on a point of its long-period[3] supply curve when three related conditions are fulfilled:

i)   factors of production are made available to the industry in the required quantity and quality (see, e.g., *Principles*, p. 497 and the masterly example concerning the fishing trade, pp. 369-372);

ii)   their organisation is the cheapest allowed by technical knowledge (see, in particular, *Principles*, Book IV, Ch.s VIII-XIII, Book V, Ch. VIII);

iii)   the price at which *that* regular flow can be sold "just suffices to induce capitalists to invest their material capital, and workers of all grades to invest their personal capital in the trade" (*Principles*, p. 497).

The *first condition* implies that, as the industry is expanded, its additional inputs are either newly produced or drawn from other industries. This condition, which recurred in all of his illustrations, normally involved a variation in a series of input prices and "other" product prices. With reference to inputs producible *ad libitum*, he stressed the possibility that their price may fall[4]. With reference to raw materials and inputs in fixed supply, a possible price rise was stressed[5].

---

[3] It cannot be seriously denied that Marshall developed his analysis of *long-period* supply much more fully than his short-period one. This prominence was ceaselessly stressed in Book V of his *Principles* (cf. p. 403 in particular; see also pp. 377, 380, 382, 464), and it was plainly accepted by commentators. The MSS contain a distinction between four different lengths of time and emphasise the relevance of the longer ones; then the PTDV concentrated *entirely* on long (or very long) periods of time. For the purpose of this paper, then, we shall concentrate on Marshall's long-period curve.

[4] For instance, in response to a permanent expansion of the fishing trade, "the industries connected with building boats, making nets, etc. being now on a larger scale would be organized more thoroughly and economically" (*Principles*, p. 371). See also PTDV, pp. 5-6.

[5] See e.g. MSS, pp. 138, 141; PTDV, p. 16; *Principles*, p. 415.

Price interrelatedness, whether due to the use of produced means of production or to the use of common factors by different industries, attracted Marshall's interest from the very beginning[6] and was to be analysed at some length in Book V, Chapter VI, as well as in the mathematical appendix of the *Principles*; at the beginning of the next Chapter, he remarked that

> These inter-relations can be and must be ignored in rapid and popular discussions of the business affairs of the world. But no study that makes any claim to thoroughness can escape from a close investigation of them. This requires many things to be borne in mind at the same time. (Marshall, 1920, p. 403-4)

It is worth stressing, then, that the *coeteris paribus* assumption which underlies Marshall's partial equilibrium analysis of supply was indeed very flexible. There was an element of judgement in the choice of the variations to be considered as the direct consequence of a change in one industry's output, and those to be embedded in *coeteris paribus*. The modern conception of partial equilibrium is more precise but much narrower and betrays Marshall by oversimplifying his argument!

Turning to the *second condition*, Marshall had a clear concept of cost minimisation *by the individual firm*[7]. Of course, cost minimising methods of production were subject to given conditions of the industry: as the industry is expanded or contracted, the firm's minimum costs changed in relation to any external economies/diseconomies[8]. The logic of the supply

---

[6] See in particular Marshall's "A note on joint and composite demand", in MSS, pp. 160-164.

[7] Very explicit statements can be found in Marshall, 1920, p. 404.

[8] In this respect, the controversial notion of the "representative firm" is central. His taste for economic facts compelled Marshall to consider industries which comprise *different* firms: hence the necessity of some average, abstract representation (incidentally, in the PTDV and in the MSS, Marshall argued precisely in terms of an *average* of actual firms). Since, however, the long-run variety of firms was gradually given less and less emphasis in later literature on competitive supply, and it is now often accepted that a long-period supply curve represents an industry formed by *identical* firms, we can put Marshall (as well as later authors) "on a par" with the current treatment of the subject, ignoring any *specific* problems arising from the variety of firms. The reader

curve led Marshall, especially in the PTDV, to underestimate the "internal" economies and to lay the main emphasis on the "external" ones. He thought that, if the economies at the level of the individual firm were important, "it would be reasonable to object that the introduction of economies into the process of manufacture does not depend directly and in the main on the magnitude of the total amount of the commodity produced' (PTDV, p. 7). This was a source of criticism and controversy, as we know. The potential conflict between Marshall's supply curve and internal economies was to be solved in later literature by assuming that the individual firm has either strictly constant returns to scale (e.g. Barone, 1894) or locally constant returns to scale, at the minimum point of U-shaped average cost curves (as in the case of Pigou, 1928 and Viner, 1953 [1931]).

This leads us to the *third condition* which defined Marshall's long-period supply curve: the absence of pure profits. Marshall's "value in use to the seller" of the MSS, his "remunerative price" of the PTDV and his "normal (long-period) supply price" of the *Principles* are equal to the expenses of production per unit, allowance being made for ordinary profits (interest) (cf. MSS, p. 138; PTDV, pp. 3 and 10; *Principles*, pp. 342-343, 411-412, 416, 419).

In the PTDV and in the first edition of the *Principles*, the whole argument is carried out in terms of the average cost per unit. The increasing stress laid in successive editions on marginal costs[9] reinforced the need to assume that the firms have either strict constant returns to scale or U-shaped average cost curves and that they are in a Wicksellian "full equilibrium".

_____

is therefore asked to remember that by a "firm" we do not mean here Marshall's representative firm or even an average firm proper, but the firm whose replication forms the industry.

[9] See the "Preface" to the 8th edition of the *Principles*, p. xvi. In the English version of his *Principii di economia pura*, Maffeo Pantaleoni claimed that the supply price should be equal to marginal, rather than average cost: "in Prof. Marshall's diagrams the supply curve is a curve of expenses per unit in function of quantity produced. It may seem doubtful whether it is convenient to consider the intersection of such a curve with the demand curve. (…) For the uses to which it is put by Prof. Marshall, a curve of marginal expenses, or marginal cost in money might be preferable" (Pantaleoni, 1898, p. 192, n 1).

All these conditions being satisfied, Marshall's analysis concentrates, of course, on the *direction of change* in the supply price as the industry output is hypothetically changed[10]. Once again, it is useful to set the mature construction of the *Principles* against the background of the more concise early versions. In the early formulation, "the only law to which the Supply curve must in all cases conform" is that it "cannot bend backwards" (PTDV, Proposition XVIII, p. 6. See also MSS, p. 145), because there cannot be more than one "remunerative price" for the same amount. With this logical restriction, Marshall considers long-period curves of all possible slopes – everywhere increasing, everywhere decreasing, alternately increasing and decreasing– as potential descriptions of market circumstances. Even though the *Principles* contains a much more articulated analysis, the conclusion is very much in the same spirit as the fundamental assumption of the early versions: "The two tendencies towards increasing and diminishing return press constantly against one another" (*Principles*, pp. 318-9). The supply curve for a certain industry, then, was simply designed to *illustrate* the outcome of one possible *set* of market circumstances of supply. Many different factors were at work *simultaneously*: as the demand for a commodity rises permanently and aggregate production increases, some raw materials are supplied to the industry at an increasing price (e.g. MSS, pp. 138 and 141, PTDV, p. 16), some machinery is supplied at a falling price (e.g. PTDV, pp. 5-6; *Principles*, p. 371), some kinds of labour are made available at a rising wage (e.g. MSS, p. 138; PTDV, p. 16), some technological and organisational economies dependent upon the industry's scale are introduced (e.g. MSS, p. 140, PTDV, pp. 9-10; *Principles*, Book

---

[10] According to Pantaleoni, the *specific* contribution of Marshall's supply and demand curves consisted precisely in the *relationship* between the industry's output and the long-period price. The conception of a long-period price at a given level of production was the same as Ricardo's, in Pantaleoni's interpretation. As he put it: "Ricardo's theorem, according to which, under conditions of perfectly free competition, commodities susceptible of reproduction are exchanged in accordance with the ratio of the costs, necessitates our considering the cost of production as the index of the available amount of every commodity. This doctrine is summed up in (…) elegant theorems of Professor Marshall" (Pantaleoni, 1898, p. 190).

IV, Ch. X). The slope of the Marshallian long-period supply curve, both in its early and its mature formulations, reflects all these factors simultaneously.

Marshall was very thrifty in his use of formalisations and this was particularly true of his theory of long-period supply:

> When making lists of supply prices (supply schedules) for long periods (…), we set down a diminished supply price against an increased amount of the flow of the goods; meaning thereby that a flow of that increased amount will in the course of time be supplied profitably at that lower price, to meet a fairly steady corresponding demand. We exclude from view any economies that may result from substantive new inventions; but we include those which may be expected to arise naturally out of adaptations of existing ideas. (…) *But such notions must be taken broadly*. The attempt to make them precise over-reaches our strength. And if we include in our account nearly all the conditions of real life, the problem is too heavy to be handled; if we select a few, then long-drawn-out and subtle reasonings with regard to them become scientific toys rather then engines for practical work. (*Principles*, pp. 460-61; emphasis added)

Nevertheless, in the Mathematical Appendix he did present his supply curve "for long periods" in terms of an *equation*, under the simplifying assumption that "all the difficulties connected with the law of increasing return" (*Principles*, Mathematical Appendix, Note XIV bis, p. 852) were ignored[11]. This is called the "supply equation" (*Principles*, p. 852). With some change in notation, it reads:

$$p = \mathbf{w}(Q)\cdot\mathbf{a}(Q). \qquad (1)$$

A commodity whose price is *p* and whose *aggregate* output is *Q* is assumed to be produced by means of a series of factors whose amounts per unit of product are the terms of vector $\mathbf{a}$, and whose *supply prices*[12] are the terms of vector $\mathbf{w}$. In Marshall's equation, $\mathbf{w}$ and $\mathbf{a}$ are

---

[11] "To adopt any other course would lead us to mathematical complexities, which though perhaps not without their use, would be unsuitable for a treatise of this kind". (Marshall, 1920, p. 852)

[12] Each function $w_i = w_i\big(a_i(Q)\cdot Q\big)$ is itself called a "supply equation".

vector *functions* of *Q*: in fact, the factor supply prices generally change along with the amount which is called forth by the industry (and equal to $a_i \cdot Q$, $i = 1, 2, \cdots, n$), in accordance with Marshall's general theory of supply. Also the terms of **a** "are generally not constant but functions of [*Q*]"(*Principles*, p. 852. In Marshall's MSS there is a "supply equation" which is similar, except for an assumption of constant coefficients of productions: see MSS, p. 160 and the equation on p. 162). Marshall omits any explicit reference to the relation between factor use by the firm and factor price. It should be surely taken for granted, however, that by "**a**" Marshall means here the *cost minimising* vector: the formula allows for such an interpretation *as it is*, because factor prices are, in turn, functions of *Q*, so that the functions $\mathbf{a}(Q)$ can be assumed to involve *both* an external economies/diseconomies effect and a substitution effect. Marshall's supply equation reflects all the elements discussed above: zero pure profit in the firms of the industry *and* in the economic processes leading to the supply of its factors of production, full adaptation of these factors to the needs of the industry, cost minimisation. Marshall regarded his mathematical analyses of the supply curve "as sketches rather than complete studies" (*Principles*, p. 852). The subsequent attempts at a formalisation by interpreters and commentators can be seen as qualifications, extensions, restrictions, or clarifications of this simple and fundamental formula.

3. Formalising the "full equilibrium" supply curve: I) Barone

The first notable contribution was that of Barone (1992)[1894], who proposed a formalisation of Marshall's curves of supply and demand in the framework of a Walrasian model of general equilibrium. In so doing, he argued that Marshall's *product* supply curve may *not* be well defined. This is the first analytical criticism of the "supply curve" and deserves a detailed discussion.

Marshall's formula is simplified by Barone who assumes a constant vector **a** (as Marshall did some 20 years before!) and concentrates on the functions $\mathbf{w}(Q)$. Two main cases should be distinguished, according to Barone. The first is when all factors are *specific to the industry*: he shows that in this case Marshall's equation is valid as it stands[13], and only requires a detailed analysis of the formation of factor supply prices. The second and more problematic case is when some factors are used in common with other industries (factor "composite demand", in Marshall's terminology). This circumstance required *ipso facto* the simultaneous consideration of at least *two* industries[14], so that Barone developed a model with two commodities, which we label α, β^{15} (each produced by a single-product firm in "full equilibrium", which operates as if it were in free competition) and two factors, labelled *l*, *k* (both owned by two subjects, labelled 1, 2): it should be noted from the outset, however, that he did *not* pioneer the familiar HOS model with fixed factor supplies, but developed a model with *variable* factors. For interpretation in the framework of fixed factor supplies, we have to wait until the early 1930s, as will be seen.

Factor supply is derived "according to Walras's equations for individual maximum utility" (Barone, 1992, p. 18). Denoting by L_i, K_i, $i = 1, 2$ the supply of each factor by each subject, and by w_l, w_r the prices per unit of the two factors in terms of money ("a means of transaction", p. 18) we have:

---

[13] In Barone (1992), the equation at the bottom of p. 34 is nothing other than Marshall's equation, in different notation. In a footnote he adds that "sometimes [Marshall] explicitly says that he takes it [the product supply curve] to be the particular case in which there is no *composite demand* for the factor services that are required to produce that product; sometimes he does not mention this assumption. However, an astute reader should always be prepared to make this assumption" (Barone, 1992, p. 36).

[14] The need to consider a group of industries and simultaneous equilibrium in the case of composite demand had been envisaged by Marshall in the MSS. See the equations on p. 164.

[15] In what follows, we have slightly changed Barone's notation, for reasons of homogeneity.

$$\frac{\psi_{il}(L_i)}{w_l} = \frac{\psi_{ik}(K_i)}{w_k} = g_i \qquad (2)$$

where ψ is the *marginal cost* for a subject to provide a certain amount of a factor, and g_i is the marginal utility of money[16] for subject i. The crucial assumption made by Barone is that g_i be determined at an initial point of Walrasian equilibrium and that it *remains constant* in the neighbourhood of this point[17]. We can thus differentiate (2) holding g_i constant, obtaining[18]:

$$dL_i = \frac{\psi_{il}(L_i)}{w_l\psi'_{il}(L_i)}dw_l \text{ and } dK_i = \frac{\psi_{ik}(K_i)}{w_k\psi'_{ik}(K_i)}dw_k, \; i = 1, 2.$$

Summing across subjects, and setting $L = L_1 + L_2$; $K = K_1 + K_2$, we have

$$dL = c_l dw_l \; ; \; dK = c_k dw_k$$

where c_l and c_k are calculated in an obvious way and express, in Barone's interpretation, the (inverted) slopes of the Marshallian curves of *factor* supply at the point of equilibrium. In the neighbourhood of a point of Walrasian equilibrium we then have

$$w_l = w_l(L); w_k = w_k(K) \qquad (3)$$

which expresses Marshall's "efforts and sacrifices" (per unit of factor), in terms of money. Now Barone's critique is precisely that, when factors are common to different industries, there is no one-to-one correspondence between a change in the *individual* industry's output, and the induced change in factor supply prices per unit. In fact, let the matrix $\begin{pmatrix} a_{l\alpha} & a_{l\beta} \\ a_{k\alpha} & a_{k\beta} \end{pmatrix}$

---

[16] Called the "standard of utility", by Barone.

[17] This assumption is coherent with Marshall's idea "that the ordinary demand and supply curves have no practical value except in the immediate neighbourhood of equilibrium" (*Principles*, p. 384 n.2; see also PTDV, p. 5).

[18] See Barone (1992), equation at the bottom of p. 26.

denote factor use per unit of output in the two industries. When factors are supplied in the required amounts, as Marshall assumed, we have

$$\begin{pmatrix} L \\ K \end{pmatrix} = \begin{pmatrix} a_{l\alpha} & a_{l\beta} \\ a_{k\alpha} & a_{k\beta} \end{pmatrix} \begin{pmatrix} Q_\alpha \\ Q_\beta \end{pmatrix} \qquad (4)$$

Substituting (4) into (3), we have

$$w_l = w_l \left(a_{l\alpha} Q_\alpha + a_{l\beta} Q_\beta \right); \quad w_k = w_k \left(a_{k\alpha} Q_\alpha + a_{k\beta} Q_\beta \right)$$

In full equilibrium, the *commodity* prices are such that

$$\begin{pmatrix} p_\alpha & p_\beta \end{pmatrix} = \begin{pmatrix} w_l & w_k \end{pmatrix} \begin{pmatrix} a_{l\alpha} & a_{l\beta} \\ a_{k\alpha} & a_{k\beta} \end{pmatrix}$$

or[19]

$$\begin{aligned} p_\alpha &= a_{l\alpha} w_l \left(a_{l\alpha} Q_\alpha + a_{l\beta} Q_\beta \right) + a_{k\alpha} w_k \left(a_{k\alpha} Q_\alpha + a_{k\beta} Q_\beta \right) \\ p_\beta &= a_{l\beta} w_l \left(a_{l\alpha} Q_\alpha + a_{l\beta} Q_\beta \right) + a_{k\beta} w_k \left(a_{k\alpha} Q_\alpha + a_{k\beta} Q_\beta \right) \end{aligned} \qquad (5)$$

Let us compare Barone's eqs. (5) with Marshall's eq. (1). It is clear that if, say, $a_{l\beta} = 0 = a_{k\alpha}$ (specific factors), then eq. (5) reduces to

$$\begin{aligned} p_\alpha &= a_{l\alpha} w_l \left(a_{l\alpha} Q_\alpha \right) \\ p_\beta &= a_{k\beta} w_k \left(a_{k\beta} Q_\beta \right) \end{aligned} \qquad (5')$$

Each equation in (5') is quite obviously a special case of (1). By contrast, if the two factors are used by both industries, neither commodity has a supply *curve*. In Barone's words:

in the case of two products, each has an *area* of supply, as [p_α and p_β] are functions of two variables [Q_α and Q_β]. (…) In a case where there are more than two products, a spatial representation is no longer possible. (Barone, 1992, p. 32; emphasis in original)

According to Barone, to draw Marshall's supply curve for, say, commodity α, it is necessary to *assume* that Q_β remains constant. This, however, "could be a serious mistake" (p. 33)

---

[19] See Barone (1992), equations in the middle of p. 32.

because Q_β can hardly be assumed to be independent of *both* p_α and p_β, which change along with Q_α.

Barone's interpretation of Marshall's supply curve is to some extent biased by the author's liking for the general equilibrium theory of his time. In particular, the assumption of constant returns to scale and, even worse, of constant coefficients, prevents his formalisation from considering the technological aspects of the "laws of return" and the problems of cost minimisation at the level of the individual firm; correspondingly, the whole emphasis is on the psychological costs of providing additional amounts of factor services. One general aspect is worth stressing. Barone takes it for granted that in Marshall's theory of supply a change in one output – determined, say, by a change in demand – is always accompanied by a change in many factor and product prices: we have a sort of parametric curve in the space and drawing the supply curve of a commodity (whenever it is well defined) amounts just to looking at this curve in the space from the partial point of view of the "own" price and quantity of that commodity. This conception was agreed upon by later interpreters and commentators, to whom we now turn.

4. <u>Formalising the "full equilibrium" supply curve: II) Pigou, Viner, Harrod, Robinson</u>

The so-called "cost controversy" led to a rethinking of the micro-foundations of Marshall's theory of supply. In the late '20s and early '30s, therefore, some new formalisations of Marshall's supply curve were proposed, and the "laws of return", neglected by Barone, were reconsidered against the background of an improved theory of the competitive firm.

A first important contribution is that of Pigou (1928)[20]. We shall stick to our convention of ignoring the long-period variety of firms, so that Pigou's argument will be referred here to industries formed by identical firms[21].

In strict accordance with Marshall, he defines "the normal supply price of any quantity of output as the price which will *just* suffice to call out a regular flow of that commodity when the industry under review is *fully adapted* to producing that quantity" (Pigou, 1928, p. 238; emphasis added); this normal supply price is calculated assuming that the price of "materials, machinery and so on (…) will vary with variations in the demand for them on the part of the industry as a whole" (Pigou, 1928, p. 241): his formalisation, then, is quite explicitly concerned with a "full equilibrium" view of supply, like Barone's. Differently from Barone[22], however, he assumed that input use per unit of output in the firm *changes* along with the output of the *firm* itself as well as with the output of the industry as a whole. His analysis was indeed inspired by the need to coordinate Marshall's internal and external economies/diseconomies with a theory of the competitive firm. For our purposes (and using our notation), we can summarise his argument as follows.

Let the cost function of the firm be[23]:

$$C = C(q,Q)$$

where q is the output of the *firm*. Pigou did not consider *input prices* as explicit arguments of the cost function, but certainly "C" is to be interpreted as a *minimum* cost for any given pair (q, Q) and given input prices. The average cost of the firm is clearly $C(q, Q)/q$; if the firm is

---

[20] See also Pigou (1929), Chapter XI of Part II and Appendix III.

[21] In order to avoid the controversies which surrounded Marshall's concept of the "representative firm", Pigou actually referred all of the argument to what he calls the "equilibrium firm" –a conception which made no impact in later literature!

[22] Unfortunately Pigou did not quote Barone; he referred directly to Marshall's *Principles*, without any explicit consideration of other authors (with the exception of Chapman and Ashton).

[23] See Pigou, 1928, p. 241

sufficiently small with respect to the industry, its marginal cost can be defined as $\partial C(q, Q)/\partial q$.

In "full equilibrium", the firm maximises its profit, *and* the industry produces a regular flow of product, with no net entries. *Both* conditions are satisfied when[24]

$$\frac{\partial C(q, Q)}{\partial q} = p = \frac{C(q, Q)}{q} \qquad (6)$$

Pigou then discusses the obvious taxonomy of possible cases and of possible corresponding equilibria of the firm, noting, of course, that it is only in the case of U-shaped curves that "there is a single point of stable equilibrium" (Pigou, 1928, p. 248), at the bottom of the average cost curve, where q and p are determined, at any given Q (and the implicit input prices).

We may evaluate the vector of input use per unit, **a**, at the bottom of the average cost curve: clearly $\mathbf{a} = \mathbf{a}(Q)$ if there are external economies/diseconomies. Moreover one should also consider all potential input prices in relation to the industry's output, so that we have $\mathbf{w} = \mathbf{w}(Q)$. At the bottom of the average cost curve, therefore, Marshall's eq. (1) holds as it stands[25]. As the *industry's* output changes, p is bound to change (as well as q). In fact, differentiating (1), under the condition that (6) be satisfied, we get

$$\frac{dp}{dQ} = \mathbf{a}\mathbf{w}_Q + \mathbf{a}_Q\mathbf{w},$$

where $\mathbf{w}_Q$ and $\mathbf{a}_Q$ are vectors of derivatives. Pigou's discussion of the overall sign of dp/dQ distinguishes, then, between the factors affecting $\mathbf{w}_Q$ and those affecting $\mathbf{a}_Q$.

---

[24] See Pigou, 1928, pp. 244-5.

[25] See Pigou, 1928, p. 251.

It should be clear, first of all, that any possible argument on the sign structure of $\mathbf{aw}_Q$ encourages the theorist to think about the economic system as a whole. This is plainly recognised by Pigou. With reference to produced means of production he points out that:

> To determine whether in fact the price of materials, machinery and so on supplied to an industry by others will rise, fall or remain constant when the output of that industry increases, we should need to step outside the industry primarily under review and investigate the conditions of production in the others. (Pigou, 1928, pp. 249-250)

In the case of primary inputs, he remarks that they "are diverted to our industry" from other industries: this transfer involves input price changes, and "they may be either positive or negative", except when "the industry does not employ more than a very small portion of the total supply of any factor, and in this latter case they are unaltered by changes in the output of our industry" (Pigou, 1928, p. 251). (The "smallness" argument is mentioned very cursorily and seems not to have been taken too seriously by Pigou, who never proposed the shortcut of simply *assuming* $\mathbf{w}_Q = \mathbf{0}$). This readiness to "step outside the industry primarily under review" is very "Marshallian", as we remarked in § 2. It must be stressed, however, that Pigou did not go very far in this direction – certainly much *less* far than Barone – and he was apparently satisfied with saying that there are no theoretical restrictions on the sign structure of the elements of $\mathbf{aw}_Q$ or on their sum (this, too, can be claimed to be "Marshallian"!).

Turning to the sign of $\mathbf{a}_Q\mathbf{w}$, Pigou made, by contrast, some definite and bold statements: $\mathbf{a}_Q\mathbf{w} > 0$ is considered *impossible* because "apart from changes in the prices of the several factors, no arrangements possible before are *excluded*" (Pigou, 1928, p. 253), and $\mathbf{a}_Q\mathbf{w} < 0$ is "materially possible" (p. 252), because "as (…) the general demand grows, it becomes more and more worth while for firms to specialise" (p. 252). Here we have a conceptualisation of increasing returns due to strictly technological factors internal to the industry, but external to the firm: the assumption that the firm always operates at the bottom of its average cost curve

was seen by Pigou as a solution to Marshall's ambiguities concerning the distinction between internal and external economies: "The essential point is that an increase in the scale to which an industry is producing frequently alters – in general diminishes – the average (and marginal) costs of the equilibrium firm contained in it, whether or not it also alters its size" (p. 252). The "material" definiteness of sign ($\mathbf{a}_Q\mathbf{w}$) and the indefiniteness of sign ($\mathbf{aw}_Q$) led Pigou to propose sign ($\mathbf{a}_Q\mathbf{w}$) as an independent conception of the "laws of supply price":

> There is then no difficulty in seeing that the law of decreasing supply price, *as conceived when correction has been made for transfer elements in rates of price change*, is not merely formally possible, but is likely to be followed in practice by many manufacturing industries. (p. 252; emphasis added)

Pigou's analysis of the supply curve was further developed and clarified by the classical 1931 article of Viner, later reprinted with the addition of a supplementary note (Viner, 1953). The analysis of costs of the individual firm, in particular, is developed in more detail, distinguishing between the short run and the long: this development led to the well-known Wong-Viner envelope theorem, by which the long-period average cost curve of the firm[26] is shown to be formed by the envelope of infinitely many short-period curves[27]. With reference to the long period, Viner adopted, like his predecessors, a "full equilibrium" view of the firm: "for long-period equilibrium not only must marginal cost of output from existent plant equal price *for each individual producer*, but it must also equal average cost". (Viner, 1953, p. 206).

---

[26] Viner "maintained, explicitly or implicitly, that under long-run static competitive equilibrium marginal costs and average costs must be uniform for all producers" (Viner, 1953, p. 222), any difference in efficiency being "compensated by differential rates of compensation to the factors responsible for such differences" (p. 201). Since, however, he made almost nothing of such long-run differences, we may safely refer his argument to the case of identical firms, as we have done for the other authors.

[27] Curiously enough, the original article contained the famous error originating precisely in the *failure* to draw the envelope of short-run cost curves properly, as Viner recognized in the 1953 *Supplementary note* (Viner, 1953, p. 227).

Like Pigou, in fact, he considered the equilibrium of the industry as a condition for the equilibrium of the firm[28]. At the bottom of the average cost curve, where the firm's plant and output have a critical scale, we can once again calculate, then, the long-run technical coefficients of the firm, as functions of the *industry's* output and present also Viner's analysis of the supply curve in terms of Marshall's equation (1).

His preliminary taxonomy can easily be reproduced by differentiating (1). The negative (positive) terms of $\mathbf{a}_Q$ are called "external *technological* economies (diseconomies)"[29], and the negative (positive) terms of $\mathbf{w}_Q$ are called "external *pecuniary* economies (diseconomies)"[30]. The aggregate change, $\mathbf{aw}_Q + \mathbf{a}_Q\mathbf{w}$, is called "*net* external economies (or diseconomies) of large production" according to whether it is negative or positive[31].

Despite the symmetry of this taxonomy, however, he did not consider each case of external effects to be equally plausible. "Illustrations of technological external economies are difficult to find" (Viner, 1953, p. 217)[32]; "External technological diseconomies, or increasing technical coefficients of production as output of the industry as a whole is increased, can be theoretically conceived, but it is hard to find convincing illustrations" (Viner, 1953, p. 221). His main emphasis, then, was on what he called *pecuniary* effects. He insists –perhaps more strongly than any other interpreter of Marshall- that while the single "concern is not of

---

[28] "The procedure which will be followed, will be to begin in each case with the mode of adjustment of a particular concern to the given market situation *when the industry as a whole is supposed to be in stable equilibrium*" (p. 200; emphasis added).

[29] See Viner, 1953, pp. 217-8 and 221.

[30] See Viner, 1953, pp. 218-19 and 220.

[31] Viner presented a similar taxonomy with reference to the *firm's* scale of production and to *internal* economies/diseconomies (pp. 212-17). This taxonomy, however, is of secondary importance for the supply curve: in fact, in the long period, the firm always operates at the bottom of its long-period average cost curve and by the Wong-Viner theorem both the plant and the output are uniquely determined.

[32] He then refers to "cross-fertilization" and "exchange of ideas" as factors which "appear to be possible sources of technological external economies resulting from an increase in the size of the industry as a whole" (p. 218).

sufficient importance to bring about any change of the prices of the factors as a result of a change in its output" (p. 201), a whole industry *can be* of a sufficient importance. This is particularly evident in the case of specific factors: if "a given industry is already utilising all of the supply available at any price of a necessary factor of production (…) it is impossible (…) to adhere to the assumption that the prices of all the factors remain the same" (Viner, 1953, pp. 207-8). Once again, however, the different conceivable cases were not equally plausible, according to Viner, and he stressed pecuniary *diseconomies* much more than pecuniary economies. Like Pigou (and Marshall), he recognised, of course, that "if industry *A* purchases materials in greater quantity, their price may fall because industry *B* can then produce them at lower unit cost" (Viner, 1953, p. 218). But he particularly insisted on the fact that

> Although it has not ordinarily been given consideration, the case of net external diseconomies of large production is of *indisputable practical importance*. Pecuniary diseconomies of this kind will *always* tend to result from the expansion of output of an industry because the increased purchases of primary factors and materials which this entails must tend to raise their unit prices. (Viner, 1953, p. 220; emphases added)

This case for a *general* tendency to pecuniary diseconomies deserves more detailed attention. Not only did Viner devote all of his "Supplementary note" to this single aspect, which he evidently considered after twenty years to be particularly important, but many other authors, such as Harrod (1930), Hicks (1934), and Robinson (1953) [1941] scrutinised it from a *logical* point of view. Viner's pecuniary diseconomies were, in fact, of crucial importance for studies in the Marshallian tradition, because the old distinction between a rising supply price in agriculture and a falling one in manufacture had long been at a dead end, and the new argument seemed to provide a firmer and wider foundation for a generalised rising supply curve, even ignoring diminishing physical returns.

The logical core of the argument put forward by the above mentioned authors consists in the fact that an industry can use a larger amount of factors *in fixed supply* only if the latter are *released* by other industries. But normally they are not released in the exact proportions needed by the expanding industry: there will be a change in relative input prices whose necessary effect will be to raise the supply price of the commodity in question *relative* to "other" commodities (as well as to economise the relatively scarce factor).

An early albeit somewhat vague hint at this mechanism can be found in Schumpeter (1928), and a much more precise version of it in the final part of Harrod (1930):

> Equilibrium is stable, provided that supply price –the price of "willingness to sell"- is an increasing function of quantity of product. This condition rests on the fundamental fact that the extending production by any given industry means withdrawing quantities of factors of production from increasingly "important" other uses. (Schumpeter, 1928, p. 365)

> Different industries mix the factors of production in different proportions. (…) Since, *ex hypothesi*, [the industry under consideration] uses more than an average amount of [factor] *A*, it can only get an increase of its output at an enhanced price. (…) Consequently the supply price of the product will on this account alone rise in response to the rise of demand for it. (Harrod, 1930, p. 240)

For a complete account, however, we have to wait until Hicks (1983) [1934] and then Viner (1953):

> If the demand for [commodity] *X* rises relatively to the demand for [commodity] *Y*, this will now involve a change in the ratio of factor prices. The price of that factor which is more used in *X* will rise. But since the price of *Y* is to remain unchanged, this means that the price of that factor more used in *Y* must fall. But it has to fall only so far as to keep the *Y* combination of factors unchanged in total price (per unit of product) and there is (*ex hypothesi*) less of the factor, whose price has fallen, in the *X* combination than in the *Y* combination. The *X* combination, containing more of the factor whose

price has risen, and less of the factor whose price has fallen, must therefore rise in price. (Hicks, 1983, p. 240)

Let us assume that in an otherwise stable economy a shift of wants occurs from other commodities to cloth, with a consequent expansion of the output of cloth. Except by coincidence, (…) the cloth industry will be using the various "factors" (…) in proportions somewhat different from those in which the economy as a whole , and the contracting section of it, uses them. As the cloth industry expands, therefore, and bids for more factors, the contracting industries will not, at the prevailing prices for the factors, be releasing factors in the same proportions in which the cloth industry is trying to acquire them (…). There will consequently occur a realignment of the prices of the factors, with those used relatively heavily by the cloth industry rising in price and those used relatively lightly by it falling in price. (Viner, 1953, p. 228)

Some critical remarks are in order here. The argument is crystal clear when referred to the case of two commodities and expressed in terms of the direction of change of their *relative price*. But Hicks and Viner used this argument to support the case for a rising supply price *in terms of money* and the connection between the two "results" is far from obvious.

This critical point has been acutely noted by Joan Robinson:

A relative increase in demand for alpha therefore entails an increase in the total demand for those factors which it employs in more than the average proportions and a decline in demand for those factors it employs in less than the average proportions. The relative prices of factors therefore alter. We are then confronted with the question: in what terms are we to measure the resources employed in alpha? We cannot say whether or not the supply price of alpha rises with an increase in its output until we know how prices are to be reckoned. (Robinson, 1953, pp. 236-37)

Robinson offered an ingenious solution, which consisted in measuring all prices "in terms of a composite unit of resources, the factors being weighted by the proportions in which they are found in industry as a whole" (Robinson, 1953, p. 237). Following Robinson's idea, Viner

developed an example[33] with a constant money income, and he obtained rising supply prices for all commodities involved. By contrast, Hicks took the price of commodity *Y* as fixed when discussing the supply price of *X*, obtaining once again the required result.

Neither choice is very convincing, however. In fact, there can be no expansion of industry *X* without a contraction of industry *Y*: Viner's numeraire (the composite good making up the national product) involved *changing* weights (that is, changing numéraire) moving from one equilibrium to another; Hicks numéraire (commodity *Y*) involved making absolutely no sense of the supply price of commodity *Y*! More generally, the economic theorist must ask whether any result on rising or falling supply prices – including Pigou's – does not depend on the *specific* and arbitrary choice of numéraire. Since it is beyond any doubt that only relative prices matter in any microeconomic argument, the common practice of measuring the supply price in terms of money conveys the idea that any qualitative result on the slope of the supply curve will be independent of the choice of (real) numéraire. Nobody, however, has so far proved this to be the case: curiously enough, such a fundamental aspect of the supply curve still requires further analysis after so many decades[34]!

5. <u>The emergence of the conventional long-period supply curve: Kaldor, Hicks, Allen, Samuelson</u>

Marshall's supply curve, and the successive versions of it discussed so far, were essentially based on "external" effects, such as changes in input prices or changes in technological efficiency due to a varying industry output. The authors in question considered the *industry* as the primary object of analysis: cost conditions in the individual firm (i.e. costs as a function of the firm's output) and the equilibrium size of the firm were, in fact, dependent upon the

---

[33] See Viner, 1953, table on p. 231.

[34] This point has been developed analytically in Opocher and Steedman (2005).

industry's output. In this sense, their very conception of the firm was dependent on their conception of the industry.

In the mid-1930s there was a drastic change in the way of thinking about the firm in relation to the industry. The intellectual centre of this "revolution" was the *London School of Economics* and its theoretical background was the Anglosaxon reception of the Walrasian paradigm of general equilibrium. Kaldor (1934) explicitly attacked the Marshallian tradition in the very pages of the *Economic Journal* where Marshall's theory of supply had been presented, criticised and developed:

> Since Marshall's time the analysis of the equilibrium of the firm has been carried to a much higher stage of refinement. In one respect, however, later constructions suffer from the same deficiency as Marshall's. (…) Explicitly or implicitly, the equilibrium of the "firm" is made dependent upon the equilibrium of the "industry" rather than the other way round. (Kaldor, 1934, p. 63)

Quite to the contrary, Kaldor says, it was

> necessary to analyse the conditions of equilibrium for the individual firms *before* any postulates were made about the supply-function of an industry. (…) Only then can we *derive* those supply curves of various shapes which the simple two-dimensional diagram at once suggests to the mind. (Kaldor, 1934, p. 61; emphasis added)

In this "bottom-up" approach, the industry's supply curve is to be formed on the basis of individual curves, each defined *prior* to the aggregate[35]. Under perfect competition, the individual firm is assumed to be a price taker and its individual supply curve is meant to represent the optimum response to price stimuli. In the words of Kaldor, "a definite amount of a commodity will be offered by each producing unit in response to any price". Here we have a proposed adaptation of the Walrasian conception of choice in relation to prices. It goes without saying, of course, that Walras did *not* distinguish neatly between the firm and the

---

[35] See also Robbins (1934), pp. 6-11.

industry[36], and that he tended to consider supply as *infinitely* elastic with respect to price[37]: Kaldor pleaded for *an extension and adaptation of general ideas* of price-taking behaviour to the specific field of product supply. This proposed micro-foundation of the supply curve was naturally extraneous to all kinds of external effects[38] and in particular to input price changes: the object of study became the reaction of individual firms and of the industry to a *parametric* change in the price of the product *alone*, taking all other prices as constant *by definition*.

The first step in this new approach was to consider supply as the *symmetrical* counterpart of demand. The Marshallian tradition was in fact accused of having led to asymmetrical curves which cannot properly be drawn on the same diagram[39]. So Allen (1938) and Hicks (1946) [1939] presented the supply curve as an extension of the theory of exchange, in which "offer" is considered as a negative demand, along Walrasian lines[40]:

> It is, however, necessary for us to go over the ground, in order to bring out a certain parallelism which exists between the case of the firm and that of the private person. It is this parallelism which will enable us to put the laws of market conduct of the firm into a similar form to that familiar to us in the other case; and ultimately to extend the theory of exchange set out in the last chapter to take account of production as well. (Hicks, 1946, p. 78)

In this extension, the determination of the equilibrium output by the firm was one and the same thing as the determination of its equilibrium inputs: the net supply and demand by the firm in all relevant markets were to be determined simultaneously, just like the net supply and

---

[36] Samuelson traces this lack of distinction back to Walras's assumption of constant returns to scale. See Samuelson (1947), p. 79.

[37] See Walras [1896] (1969), Lesson 22 and in particular Figure 24.

[38] See Kaldor, 1934, p. 66.

[39] Cf. Kaldor, 1934, p. 60; Hicks, 1934, p. 237. Offer as negative demand was also a major feature of Wicksteed's theorising. Marshall noted that the symmetry of reciprocal demand curves which he discussed in the *Pure Theory of Foreign Trade* was broken when passing to curves of supply and demand, which he discussed in the *PTDV*: see Marshall (1930), p. 2.

[40] Cf. Walras (1969), Lesson 9, and in particular p. 139.

demand of "the private person". From a formal point of view, this required one to "treat the factors as negative products" (Hicks, 1946, p. 319) and to obtain solutions for input use and for output(s) which maximise their algebraic sum. Therefore, using first and second order conditions, the analysis was directed to obtaining a mapping from a point in price space (input prices, output price(s)) to a point in quantity space (input use, output(s),), that is, in the simple case of a single product firm

$$(\mathbf{w}, p) \rightarrow (\mathbf{x}, q).$$

In order to represent this mapping (when it exists) with functions of one variable, all other variables have to be taken as constant *by assumption*. If all input prices are assumed constant, we end up with output *and input use* as functions of the output price.

In the (q, p) space, this is the supply curve *of the firm*[41]. This approach led Hicks and Allen to formalise the theory of supply (by the price-taking firm) using the same categories as used in their theory of demand: in relation to their response to the change in one price, inputs and outputs are classified as complements or substitutes; in relation to their response to a change in output (i.e. a movement along the supply curve), inputs are classified as regressive or not regressive[42] (later to be called inferior or normal).

Even though Hicks and Allen obtained their results on the basis of the production function (or, rather, of an implicit transformation function to allow for joint production) it is convenient for us to express the same results on the basis of the cost function, $C(q, \mathbf{w})$,

---

[41] We note in passing that if the numéraire of the output price is a composite of inputs, the same curve (except for scale) is obtained irrespective of the weight given to each input, for the simple reason that all input prices are constant along the supply curve: no numéraire problem of the kind referred to in the previous section can possibly arise in this context.

[42] See Hicks, 1948, Chapter VII.

adopting the practice introduced by Samuelson (1947) [43]. The supply curve can be expressed as the solution for q, of the system:

$$p \leq \frac{\partial C}{\partial q}(q, \overline{\mathbf{w}}), \text{ with equality if } q > 0 ; \qquad (7)$$

$$\frac{\partial^2 C}{\partial q^2}(q, \overline{\mathbf{w}}) \geq 0 ; \qquad (8)$$

$$pq - C(q, \overline{\mathbf{w}}) \geq -C(0, \overline{\mathbf{w}}), \qquad (9)$$

where p is taken as a variable parameter, and $C(0, \overline{\mathbf{w}})$ is the fixed cost. (As usual, the vector $C_{\mathbf{w}}$ measures input use at any point of the supply curve).

Equations (7)-(9) define quite naturally the *short-run* supply curve of the firm, where, by definition, $C(0, \overline{\mathbf{w}}) > 0$, and (9) can be satisfied as an inequality: as price rises, both output and profits (positive or negative) rise. By the very logic of profit maximisation, then, the short-run supply curve of the firm slopes upward, quite independently of any empirical consideration. The results obtained for the firm immediately generalise to the *industry's* short-run supply curve. In the words of Hicks:

> Strictly speaking, we only discussed in the last chapter the effect of a change in price on the demands and supplies of a single firm. Here we need the effect on a group of firms. For the most part this effect can be got by aggregating the effects on single firms, as we found we could aggregate the effects on private individuals; *so far the group must obey the same laws as the single firm.* (Hicks, 1946, p. 102; emphasis added)

On the contrary, the passage from the short-period – where the new approach found its most natural application – to the long, encountered a series of problems. In fact, Hicks goes on to ask:

---

[43] Samuelson (1947) was among the first to introduce profit and cost functions: see, in particular, pp. 76-78. A complete formal theory of cost functions for the price-taking firm was first presented by Shephard (1970).

What happens, however, if the change in prices has the effect of altering the number of firms producing a particular commodity, so that firms enter or leave the "industry"? This is a notoriously tricky matter, and it is right that we should proceed with caution. (Hicks, 1946, p. 102)

To be sure, Hicks did not think that free entry as such was a serious obstacle for the construction of an industry's supply curve[44]. Rather, he stressed some problems that his theory of the *firm* encountered when referred to the long run. According to Hicks, in the long run there is a tendency for *constant returns to scale* to prevail[45]. But with *strict* constant returns to scale and zero fixed costs, *any output* satisfies (7) and (9) when

$$p = \frac{\partial C(q, \mathbf{w})}{\partial q} = \frac{C(q, \mathbf{w})}{q} \ . \qquad (10)$$

Not only is the equilibrium of the firm *indeterminate*, but prices cannot be taken as given, independently of the firm's costs. The specific question of how the price-taking firm reacts to an *isolated* change in its output price becomes an ill-defined question. This was plainly recognised by Hicks:

It is not possible for the price of one factor (or product) to change, there being no change in the prices of all other factors and products, without upsetting equilibrium altogether. If the price of a product rises, output would become infinite; if the price of a factor rises, it will become zero. In the limiting case we are considering, our analysis threatens to break down altogether. (Hicks, 1946, p. 322)

A similar concern was expressed by Allen[46] and by Samuelson[47].

---

[44] "Nevertheless it does not appear that for our present purposes the qualifications introduced by the possibility of new firms are likely to be serious" (Hicks, 1948, p. 102).

[45] "If no such fixed [productive] opportunity exists, then there is no reason why an equal proportional increase in all factors should not enable all products to be increased in the same proportion as the factors have been increased" (Hicks, 1946, p. 322).

[46] "[I]n the long period, if all elements are variable and returns are constant (…) the prices assume particular given values and no one of them can be changed by itself without destroying equilibrium" (Allen, 1957, p. 617).

It is true that the problem of indeterminacy can be overcome assuming a long-run U-shaped average cost curve, whose relevant equilibrium point is where returns are *locally* constant, so that equations (10) hold for a *determinate* level of q. This would *not*, however, eliminate the *main* problem, which consists in the logical *impossibility* of changing one price in isolation *and* keeping the firm in equilibrium: either more than one price is assumed to change, or the firm is driven out of equilibrium without remedy. A series of interesting contributions in the 1970s[48] – which have not received the attention they deserve – tried to overcome this inconsistency by making a "less partial" equilibrium analysis of the competitive firm[49].

There are, however, further problems which apparently received no attention from Hicks or Allen or Samuelson. Let us *assume* that in response to a price change, the output of the whole industry changes in a certain measure, due partly to enlargement/contraction of existing firms, and partly to entry/exit of firms. The old problem posed by Marshall, Pigou, Viner, etc. is still there: a permanent transfer of inputs from one industry to another cannot be without effect on their prices! Moreover, via input-output relations among industries, a parametric change in one price cannot be without long-period effects on a series of other prices[50]. Finally, let us assume that at one point of the industry supply curve all firms are in equilibrium; as price changes, we move to another point of the industry's curve: but if firms are to be again in long-period equilibrium, more than one price must have been changed!

The logic of equilibrium of the firm and the industry in the face of a "price signal" – which defines partial equilibrium analysis in the Hicks-Allen-Samuelson understanding – therefore

---

[47] "If competition is "pure" in the commodity and factor markets, and the production function is homogeneous of the first order, then it is a classical fact that (…) the Hessian of the production function is singular. Therefore a regular maximum for the firm is impossible" (Samuelson, 1947, p. 78).

[48] See Silberberg (1974a) and (1974b) and the references therein quoted.

[49] Silberberg, for example, changed both one input price and the product price, so as to keep profits equal to zero, and examined the consequent changes in input use per unit of output.

[50] Taking these effects into account makes a difference to the comparative static properties of equilibrium ("full equilibrium") of the competitive firm, as Steedman (1998) has shown.

reveals very serious problems, when referred to the long period. These problems have been

partially recognised; but they certainly require further analysis.

6. Long-run supply curves in current textbooks

It is now in order to consider how reputable microeconomics textbooks have accommodated

(or not accommodated) the two traditions and their critical aspects.

Let us first consider the long-run equilibrium of the *firm*.

The commonest *definition* of a long-run equilibrium of the firm is a situation in which the

firm has adapted *all* of its inputs and its output to a "permanent" price configuration. Input

prices are given, and profits are the highest allowed by the given (variable) commodity price.

They have a zero lower bound, however, due to the absence of fixed factors. In this definition,

"the competitive firm's long-run supply function (…) is identical to Long-run Marginal Cost

(…) except that, below the minimum of Long-run Average Cost (…) [it] lies along the

vertical axis" (Hirshleifer and Glazer, 1992, p. 180). As we said, the firm moves along the

long-run marginal cost curve (rather than the short-run one) *only* "if [it] believes the change in

price is *permanent*" (Hirshleifer and Glazer, 1992, p. 181; emphasis added). Along the same

lines, Mas-Colell *et al* (1995) draw what they call the firm's "long-run supply

correspondence" (See Figures 10.F.1 (a), 10.F.2(a), 10.F.4 (a)).

There is clearly a change here in the very conception of the long-run equilibrium of the firm:

the *zero* profit condition of Marshall, Walras, Wicksell, and also of Pigou, Viner, and even of

Hicks, Allen and Samuelson is replaced by a *non-negativity* condition. This change gets rid of

all the problems which arise in "full equilibrium" analysis. It raises *other* problems, however:

why, in a perfectly competitive industry, can firms "shut down" more "easily" than new firms

can enter? And if free entry is assumed, how can a given price configuration which makes

room for pure profits be taken as "permanent"? The few attempts to tackle these questions are

rather timid. A note-worthy example is that of Stigler (1966). He argues that under

competition "*only if* the industry employed specialized resources (say, a special kind of land) … would one generally expect numbers of entrants to be unresponsive to price in the long run" (Stigler, 1966, p. 182; emphasis added). Rather, the neglect of free entry is justified by Stigler on an *empirical* ground: "the empirical evidence suggests that a large part of the increases in output of a growing industry come from the existing firms" (*ibid.)*. Apart from this notable exception, however, the above questions are simply avoided by considering the firm *as if* it did not belong to a competitive industry with free entry. The now conventional long-run supply curve of the firm is therefore something of a hybrid, a very artificial conception, which just mimics the short-run curve: the only differences between the two are that in the former the "shut down" condition replaces the "zero output" condition and that in the former the marginal cost curve is assumed to be flatter (or much flatter).

By contrast, in equally reputable textbooks, and very frequently in the *same* textbook a "full equilibrium" definition can also be found, which postulates *zero* profit[51]. In this case, there is simply *no* supply curve of the firm, because just *one* price/output pair is relevant (for given input prices). A notable example is the manual of Ferguson (1969):

> PROPOSITION. Long-run equilibrium for a firm in perfect competition occurs at the point where price equals minimum long-run average cost. At this point minimum short-run average total cost equals minimum long-run average total cost, and the short- and the long-run marginal costs are equal. The position of long-run equilibrium is characterised by a "no profit" situation – the firms have neither a pure profit nor a pure loss, only an accounting profit equal to the rate of return obtainable in other perfectly competitive industries. (Ferguson, 1969, p. 244)

The same idea is still to be found, for example, in a very recent (introductory) text, Krugman & Wells (2005, pp. 219-220).

---

[51] A very clear example of this "double" conception of the long-run equilibrium of the firm is in Henderson and Quandt (1971) p. 111 and p. 117, respectively.

Turning to the *industry*, the typical argument of textbooks which admit of a long-run supply curve of the firm, is as follows. Let there be a number of (actual or potential) *identical* firms in a certain competitive industry, each with its long-run supply curve. At any given number of firms, we can define the *industry marginal cost function*[52]. As the number of firms increases, the curve becomes flatter and flatter and in the limiting case it approaches a horizontal line, which corresponds to the shut down price for each of them. The same result can, of course, be obtained more directly, starting from a "full equilibrium" view of the firm[53]!

At this point, possible "external effects" are taken into consideration. For it may seem unreasonable to assume that the cost conditions of the individual firm are completely independent of whether the industry is very small or very large. Some reputable textbooks develop this aspect, which clearly introduces into the Walras-*LSE* approach some elements of the Marshall-Pigou-Viner approach. As Hirshleifer and Glazer put it, "the individual firm is so small that expansion of its output would not significantly drive up the prices of the resources it hires"(Hirshleifer and Glazer, 1992, p. 178), but "allowance must be made for the effect of *industry-wide* output changes upon input prices" (p. 182; emphasis added); similarly, according to Ferguson, as new firms are attracted into the industry, "the usage of resources expands and now, we assume, resource prices expand [sic] with resource usage. The cost of inputs therefore increases for the established firms as well as for the new entrants" (Ferguson, 1969, p. 246). Also Stigler (1966) distinguishes the cases of increasing and decreasing cost industries according as input prices rise or fall along with the industry's output[54].

Also with respect to the industry, then, we have two textbook versions of the long-run supply curve. One of them builds strictly on the Hicks-Allen type of partial equilibrium, with constant input prices. This version ends up with a horizontal industry supply curve: in this

---

[52] See Mas-Colell et al. (1995), p. 337; Varian (1984), p. 89.

[53] E.g. Ferguson (1969), p. 244.

[54] See Stigler, 1966, pp. 184-87.

respect, things remain as they were left by Walras in his lectures[55]! The other version, as we just said, allows for variable input prices and thereby goes beyond the Hicks-Allen partial equilibrium method. Does this version contain any progress, as compared with the "old" Pigou-Viner version? We doubt it. For consistency we must look back at the *firm*: how does the firm's output react to a combined change of output price and input prices as we move along the industry's supply curve *à la* Kaldor? Are we expecting to obtain a meaningful conception of the firm's supply curve which might serve as the microeconomic basis of the industry's curve? And in terms of what are these prices to be measured, since a money measure is inadequate when the structure of relative prices is assumed to change? No answer to these questions can be found in the textbooks, for the simple reason that no satisfactory answer has ever been provided. At *best*, we are back to the Pigou-Viner version.

7. "Marshallian" and "Walrasian" supply curves?

We have not labelled the supply curves encountered in our historical reconstruction as "Marshallian" or as "Walrasian". But since these labels commonly distinguish two different conceptualisations of the supply curve, we can no longer avoid the question of whether such labels have any clear meaning.

From a strictly historical point of view, it is perfectly in order to say that there is a long Marshallian tradition, which dates back to Marshall's early writings and, for brevity, can be said to gravitate around equation (1): the elaborations and interpretations of Barone (in 1894) of Pigou and Viner (1928-1931) and of Harrod and Hicks (1930-1934), with the further developments due to Robinson in 1941 and Viner in 1953 belong to this "tradition". In 1934 Kaldor and Robbins attacked that tradition *as a whole*, with the argument that it was based on a conception of the firm which was not independent of industry equilibrium. This marked a

---

[55] See Walras, 1969, Lesson 22, Figure 24(a), p. 262.

31

turning point in theoretical investigations on the supply curve and prompted the emergence, with the analyses carried out by Allen, Hicks and Samuelson (1938-1947) of an industry supply curve *entirely based* on profit maximisation by price-taking firms. The source of inspiration was certainly the Anglosaxon reading of Walras, and the resulting supply curves (of the firm or of the industry; referred to the short or to the long run) have every historical reason to be called "Walrasian".

Nevertheless, it must be stressed that on *analytical* grounds the two labels may be misleading. In fact, as we have seen, the so-called Marshallian tradition typically used some general equilibrium arguments, whereas the Hicks-Allen type of analysis is based on a strict definition of the partial equilibrium method: are we to conclude that the "Walrasian" curve (in the historical sense) is in fact "Marshallian" (in an analytical sense) and conversely? To avoid any such play on words, we should resist the temptation to use such simple labels, either to denote an historical origin or to denote methods. Rather, we need to recognise that in respect to supply there *is no* sharp, clearcut borderline between partial and general equilibrium methods, as Schumpeter (1963, p. 994) noted: *any* analysis takes some things as independent of that which is primarily studied, and other things as dependent upon it, so that an element of judgement is always involved as to which "indirect effects" (Marshall, *Industry and Trade*, Appendix A) may reasonably be ignored.

8. Concluding remarks

The familiar supply curves are referred to the short run or to the long, to the firm or to the industry. The analytical foundations of the *short-run* curves, whose historical roots certainly consist in the Hicks-Allen formalisation of the late 1930s, are very clear and uncontroversial. By contrast, the analytical foundations of the *long-run* curves are not self-evident and their

historical roots seem to have been lost sight of. We have proposed in this paper a historical reconstruction of these foundations and found them to be far from simple.

The "*LSE*" approach to the supply curve, prompted by the Robbins-Kaldor criticism of the whole Marshallian tradition, cannot be said to have been as successful with the long run as with the short. In particular, the micro-foundation of the industry's long-run supply curve is rather weak, because the zero-profit firm is automatically pushed out of equilibrium by a parametric change in the output price (unless a change in input prices is allowed). This determined a change in the very conception of the equilibrium of the firm: from "full equilibrium" in which the firm is assumed to have *zero* pure profits to a "conventional long-period equilibrium" in which the firm is assumed to have *non-negative* profits. The first conception was the one inherited from Marshall, Walras, Wicksell etc. and had never been questioned by the "Marshallian tradition"; the second conception was original. By aggregating the "conventional long-run" supply curves of an increasing number of identical firms we obtain, as the limiting case, a horizontal line, which represents no more and no less than Walras's price equation for one commodity, when input prices are taken as given. This can hardly be claimed to be a great leap forward!

In fact, the supply curve of the "Marshallian tradition" survived the spread of the *LSE* method: some textbooks, for instance, make *nothing* of the firm's long-run supply curve, and immediately consider the industry's curve, under the assumption that each firm is in "full equilibrium". This allows for "external effects", and in particular input price changes, to be considered. This survival, however, raises a number of questions: did the search for an independent conception of the firm's supply curve merely create a pseudo-problem? How are input price changes to be determined precisely? And in a context of multiple relative price changes, what is the possible meaning of an individual "price rise" when it is dependent on the choice of numéraire?

In spite of its familiar flavour and of its secular, noble history, is it not legitimate to doubt whether the evolution of the long-run supply curve was a success story and whether, as we receive it now, it is of any use in economic analysis?

References

Allen R.G.D. (1938), *Mathematical Analysis for Economists*, London, Macmillan.

Allen R.G.D. (1957), *Mathematical Economics*, London, Macmillan, Second Edition.

Barone E. (1992) [1894], "On the Analysis of Dynamic Problems". In L. L. Pasinetti (Ed.), *Italian Economic Papers*, I: 17-38. Bologna, Il Mulino/Oxford Economic Press.

Ferguson C.E. (1969), *Microeconomic Theory*, Nobleton, Irwin-Dorsey Limited. Revised edition.

Groenewegen P. (1995), *A Soaring Eagle: Alfred Marshall 1842-1924*, Edward Elgar, Cheltenham.

Harrod R.F.(1930), "Notes on Supply", *The Economic Journal*, Vol. 40, No. 158: 232-241.

Henderson M. and R.E. Quandt (1971), *Microeconomic Theory : A Mathematical Approach*, New York, McGraw-Hill.

Hicks J.R. (1983) [1934], "A Note on the Elasticity of Supply". In *Collected Essays on Economic Theory*, Vol. III: 237-45, Oxford, Basil Blackwell.

Hicks J.R. (1946), *Value and Capital*, Oxford, Clarendon Press, Second Edition.

Hirshleifer J. and A. Glazer (1992), *Price Theory and Applications*, New Jersey, Prentice-Hall.

Kaldor N. (1934), "The Equilibrium of the Firm", *The Economic Journal*, Vol. 44, No. 173: 60-76.

Krugman P. and R. Wells (2005), *Microeconomics*, New York, Worth Publishers.

Marshall A. (1919), *Industry and Trade: A Study of Industrial Technique and Business Organization; and of Their Influences on the Conditions of Various Classes and Nations*, London, Macmillan.

Marshall A. (1920), *Principles of Economics*, London, Macmillan, Eighth edition.

Marshall A. (1930), *The Pure Theory of Foreign Trade – The Pure Theory of Domestic Values*, London, The London School of Economics and Political Science, University of London.

Mas-Colell A. *et al* (1995), *Microeconomic Theory*, Oxford, Oxford University Press.

Opocher A. and Steedman I. (2005), "Long-run Rising Supply Price and the numéraire", mimeo.

Pantaleoni M. (1898) [1889], *Pure Economics*, London, Macmillan (English translation by T. Boston Bruce, based on 2nd ed. of the Italian original; 1st Italian edition: Pantaleoni M. (1889), *Principii di economia pura*. Florence: Barbera.

Pigou A.C. (Ed.) (1925), *Memorials of Alfred Marshall*, London, Macmillan.

Pigou A.C. (1928), "An Analysis of Supply", *The Economic Journal*, Vol. 38, No. 150: 238-257.

Pigou A.C. (1929), *The Economics of Welfare*, London, Macmillan, Third edition.

Robbins L. (1934), "Remarks Upon Certain Aspects of the Theory of Costs", *The Economic Journal*, Vol. 44, No. 173: 1-18.

Robinson J. (1953) [1941], "Rising Supply Price". In G. J. Stigler et K.E Boulding (eds.), *Readings in Price Theory*, London, Allen and Unwin: 233-41.

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

Schumpeter J. (1928), "The Instability of Capitalism", *The Economic Journal*, Vol. 38, No. 151: 361-86.

Schumpeter J. (1963), *History of Economic Analysis*, New York, Oxford University Press.

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

Silberberg E. (1974a), "A Revision of Comparative Statics Methodology in Economics, or, How to Do Comparative Statics on the Back of an Envelope." *Journal of Economic Theory*, 22: 159-172.

Silberberg E. (1974b), "The Theory of the Firm in "Long-run" Equilibrium." *American Economic Review*, 64: 734-41.

Steedman I. (1998), "Produced Input Use per Unit of Output", *Economics Letters*, 60: 195-99.

Stigler G.J. (1966), *The Theory of Price*, Toronto, Collier-Macmillan.

Varian H. (1984), *Microeconomic Analysis*, New York, Norton & Company, Second edition.

Viner J. (1953), "Cost Curves and Supply Curves". In G. J. Stigler et K.E Boulding (eds.), *Readings in Price Theory*, London, Allen and Unwin: 198-232.

Walras L. (1969) [1896], *Elements of Pure Economics. Or the Theory of Social Wealth*. Third Edition. Reprints of Economics Classics, August Kelley, New York.

Wicksell K. (1935) [1901], *Lectures on Political Economy*. London: Routledge.

Whitaker J. K. (1975, Edited and Introduced by), *The Early Economic Writings of Alfred Marshall, 1867-1890*, Macmillan for the Royal Economic Society, London.