4. The compatibility of scale economies with competition

The decisive step in identifying a situation of natural monopoly is to recognize that generalized scale economies lead to the survival of a single firm in the market; graphically, if the intersection between the demand curve and the average costs curve is found in the downward part of the latter, that market will be served by a single firm. The question under examination in this section is what economists thought would happen to the market structure as the average cost decreased over the entire relevant range of market demand.

It is well known that Senior divided monopolies into four kinds; he considered one kind to be due to scale economies, but he doesn't explain how it happens that economies of scale lead to monopoly; moreover he applies this feature to production in which there are also "certain exclusive facilities as a producer" such as "patent machinery" (1836: 4.51)[32]. Cournot was the first to clearly pose the problem: if the marginal cost function is diminishing, he writes, "nothing would limit the production of the article" and a monopoly would occur ([1838] 1960: 91)Albeit basing his case on different arguments, J.S. Mill reaches the same conclusion. As we have seen, he identifies many causes of cost savings in large scale production; in these cases, according to Mill, the presence of numerous small businesses is a useless waste. All this leads him to conclude that when competition could bring about only a multiplication of costs, a single firm will survive (1848: I.9.7). As we have seen in the previous section, J. S. Mill and Dupuit inspired Walras to write : "Laying a second system of water or gas pipes in a city where there is already one that could satisfy all the needs, building a second network of roads in a country where there is already one that is enough for all the communications, would be an absurd way of chasing after economies" ([1875] 1936: 211). Ely just briefly mentions the question, but hits the nail on the head: "there is great economy and convenience in the conduct of the transportation … by those operating on a vast scale … and this gives to that industry its inherent and irresistible impulse toward monopoly" (1886: 574). Like many other economists, he also takes up J.S. Mill's argument about waste, stating: "One telegraph company can send telegrams all over the country cheaper than two" (Ely 1889: 44). Hadley makes an interesting comment on the process toward equilibrium: since the railroad "is

---

[32] This is perhaps one of the reasons why Stigler writes that Senior "was wholly promiscuous in his use of the concept of monopoly" (1957:3).

not subject to the law of the diminishing return … there is … no direct limit to [the] cut-throat competition" (1886: 35); such price competition generates, in his opinion, the survival of only one firm, and therefore a monopoly. For De Viti de Marco the telephone industry tends to be a monopoly not only due to cost features, but also due to network effects.

The position taken by Marshall (1890a) in this respect is much more complex. The literature on this aspect of his thought is massive[33]. We need only recall here that in his *Principles*, Marshall cites Cournot's statement on the contradiction between economies of scale and competition, but criticizes him for following only the abstract logic of mathematics. He tries, as we know, various ways of solving what he called the "Cournot dilemma", without abandoning either internal economies or competition[34]. Like Marshall, also Pareto critically examines the theory that economies of scale necessarily lead to monopoly, and concludes that "the facts are not in accord with this theory" ([1906] 1971: 243).

A truly complete examination of the concept of natural monopoly can be found in *Principi di economia politica* by Barone, who explains it in the following terms: "If the unit cost of the product were to diminish indefinitely, as the quantity of output increases, it would be advantageous for the production of every good to be concentrated in a single firm" (1908 [1936]: 20-21). "And this can occur – he adds – when … there is … a type of firm such that, while its costs decline toward a limit, its size is enough to satisfy the entire market demand" (289)[35].

Schumpeter writes that Marshall had "all the clues that are needed for a satisfactory treatment of decreasing costs in all its meanings and aspects"; then he adds: "the only thing to wonder at is that this discussion took so long to burst into print and to present results to the scientific public at large" ([1954] 1986: 1046). In actual fact, contributions to the issues reappear only in the 1920s[36]. It is well known that incompatibility was one of the

---

[33] We just mention Stigler (1951), Schumpeter (1954), Whitaker (1990), Marchionatti (1992), Loasby (1996), Groenewegen (1999), Whitaker (2003), Hart (2004).

[34] We also know that Marshall changed his views over the years on this question (Whitaker 2003).

[35] The same concept is found in his *Principi di economia finanziaria* (Barone 1911-12: 120-121).

[36] Buchanan and Yoon (1999) believe that, after the success of the U-shaped cost curves, neoclassical authors gave up the hypothesis of generalized increasing returns because of the difficulties it created for their theory of distribution.

focuses of Sraffa's critique of Marshallian supply function[37]. His 1925 article opens with the quotation of Pantaleoni's classification of industries based on different returns; he then states the contradiction between internal economies and competitive equilibrium[38]. He returned to this in his 1926 article: "reductions in cost connected with an increase in a firm's scale of production – he writes – arising from internal economies or from the possibility of distributing the overhead charges over a larger number of product units, [are] incompatible with competitive conditions" (1926: 540). In a brief historical reconstruction of Smith's theorem on increasing returns, Stigler (1951) does not mention Sraffa, while he cites Knight's criticisms of Marshall's solutions for the reconciliation problem. In fact, in the first half of the 1920s, Knight dealt with the problem several times[39], reaffirming that the necessary outcome of generalized decreasing costs due to large scale production was monopoly (Marchionatti 2003: 60). We can thus ascribe to Sraffa and Knight the definitive statement on the incompatibility of increasing returns and perfect competition. Of course, in the 1920s both of them, together with many other economists inside and outside Cambridge, contributed to the well known "cost controversy", which was focused exactly on whether increasing returns and competition could be compatible or not. It's worth noting that those participating in this conference didn't mention any concrete example of industries exhibiting generalized economies of scale; moreover, the expression "natural monopoly" to identify those industries was never used. Once more, we see that the different components of the concept followed different paths.

5. The diagram

We have seen that Barone perfectly described the shape of average cost curves and the situation in which natural monopoly occurs, using both cost curves and market demand. He did it in words, and didn't draw the diagram. It was Edgeworth, in one of his four famous articles on the theory of railway rates[40], written between 1911 and 1913, who drew a diagram in which natural monopoly was perfectly represented, probably for the

---

[37] The literature devoted to Sraffa's criticisms of Marshall's theory is also massive; see the up-to-date references in Rosselli (2005).
[38] For an interpretation of Sraffa's 1925 article see Roncaglia (1991).
[39] In particular we remember Knight (1921).
[40] We have already mentioned three of these four articles in the section on economies of scale.