**Baumol on Excess Capacity**

Tamara Todorova

American University in Bulgaria

Abstract

Professor Baumol is mostly credited for his study of contestable markets as markets enjoying monopoly power which could potentially turn into competitive ones due to the lack of sunk costs and entry or exit barriers. We reveal Baumol’s view of the welfare aspects of monopolistic competition. Discussing the so called “excess capacity theorem,” formulated by Robinson and Chamberlin, and basing his analysis on Machlup’s writings, Baumol justifies competition on account of the nonexistence of excess capacity. This reveals Baumol as a true proponent of competition.

Keywords: William Baumol, monopolistic competition, excess capacity, competition

JEL codes: B21, B31, D41, D43

Professor Baumol is credited for his theory of contestable markets and the hit-and-run strategy in oligopolistic industries, but little is known about his contribution to the study of excess capacity and the welfare effects of monopolistic competition. Excess capacity is associated with production in monopolistic competition and is viewed as one of its inefficiencies. Monopolistic competition thus is seen as suboptimal to monopoly and oligopoly which enjoy market power but take advantage of the scale of operations. This is an attempt to undermine monopolistic competition as the optimal market structure in the conditions of imperfect information and positive transaction costs. Excess capacity is also widely examined by economists who have studied it from various perspectives – from dynamic optimization models to general equilibrium analyses. Baumol emphasized that there may not be excess capacity at all. In “Monopolistic competition and welfare economics” he discussed the excess capacity theorem defined by Chamberlin (1947). Building on Machlup (1952), Baumol (1964) discussed three possible outcomes of the theorem. In the narrow sense of the excess capacity theorem the typical company’s demand curve may plausibly be expected to be quite flat and excess capacity correspondingly insignificant. From the perspective of industry size and the number of firms, the excess capacity theorem suggests that the same total output would be produced more efficiently and at less cost by a smaller number of firms.

Discussing the social costs of standardization, Baumol (1964) observed that if the number of firms in the industry is reduced, the variety of products available to consumers must fall. The resulting saving in resources is then to be considered a net gain or a loss and should be judged on a case-by-case basis. Baumol (1964) maintains that excess capacity represents a real social cost only if the variety consumers enjoy compensates them adequately for the additional costs it imposes. This paper investigates the concept of excess capacity, as introduced by Chamberlin (1933, 1947) and Robinson (1933). We highlight the novelty of Baumol (1964) for his early grasp of the geometrical inconsistency of this concept. To contextualize Baumol’s ideas we will contrast the views by Machlup (1939), Schumpeter (1939), Nichol (1934), Harrod (1952), Archibald (1961), Demsetz (1959), etc. We analyze Baumol’s substantive contribution to welfare economics in the cost-benefit analysis of monopolistic competition as the most common type of competition.

1. Introduction

Baumol is mostly credited for his theory of contestable markets and the hit-and-run strategy in oligopolistic industries. In a series of papers Baumol (1982) and a group of co-authors, Baumol, Panzar and Willig (1982, 1983, 1986), Bailey and Baumol (1984), Baumol and Willig (1986), demonstrated that the competitive outcome could be achieved even in industries which exhibited market power such as monopoly and oligopoly. In the absence of entry barriers and sunk costs it is easy for a competitor to enter such an industry, steal the incumbent’s profits and leave unbothered. The incumbent, therefore, is forced to keep prices relatively low to deter entry by newcomers. Such an attempt leaves the incumbent with almost no profit but produces the competitive outcome, which resembles a Bertrand type of oligopoly competition. The policy implications of the research conducted by Baumol et al. (1982, 1983, 1984, 1986) were that there need not be regulation in contestable markets because in the absence of entry and exit barriers entry-limit pricing leads monopolies and oligopolies to the competitive outcome without any government intervention. Contestability provided the theoretical ground for deregulating airlines, trucking, buses, and many other transportation markets. Deregulation had already started in the transportation business. The Airline Deregulation Act of 1978 introduced much competition in the domestic U.S. aviation industry, allowing new jet carriers to enter the industry, airline managements to freely structure their route networks as they view optimal, and prices to be set competitively. The Motor Carrier Act of 1980 liberalized the trucking industry, exempting it from various types of regulation. With technological advancement in the telecommunications sector a court decision in 1984 effectively ended AT&T’s telephone monopoly, breaking it up into several firms and allowing entry by MCI Communications and Sprint Communications. Deregulation and privatization were undertaken in expectation that potential competition among the existing large firms in various industries would protect consumers and bring optimality in each industry.[[1]](#footnote-1)

The theory of contestable markets might have left many with the impression that Professor Baumol was a proponent of market power and market structures in which firms benefit from internal economies of scale and large size.[[2]](#footnote-2) What has received almost no attention is that Baumol also contributed to the study of excess capacity and the welfare effects of monopolistic competition. In a 1964 paper “Monopolistic Competition and Welfare Economics” he defended competition on account of the non-existence of excess capacity. He believed that there was room for all types of firms, both oligopolists and competitors, in the economy. Discussing his sales maximization hypothesis and the goal of firms to achieve normal profit, i.e., the rate of interest on capital and some compensation for risk, he writes:

“If this were so and if all firms were to succeed in this goal, then a mixed world of competitors and oligopolists might yet produce the ideal output, provided, of course, that it were untouched by externalities.” (Baumol, 1964, p. 47).

In Baumol’s world of contestability and sales maximization, different market structures are optimal in different situations, but each has a place in the economy. This is broadly consistent with his later theory of entrepreneurship and innovation. An analysis of his earlier views can thus help us figure out what was new to his later theoretical contributions. Baumol believed that analyzing the allocation of resources under different market structures and comparing the actual outcome to the social optimum is one of the key achievements of welfare theory. This social optimum is the so-called “ideal output” which economics has sought to find in every market structure. Welfare economists are always sensitive about the degree of competitiveness in an industry, the abuse of market power and the loss of total social surplus.

2. Baumol on excess capacity

In 1933 Chamberlin formulated the so-called excess capacity theorem according to which there is unused capacity with monopolistic competition since the firm does not produce at the minimum point of the long-run average cost curve but at a point left of it at the tangency of the average cost curve with the demand function. Both Robinson (1933) and Chamberlin (1933, 1947) discussed the downsides of monopolistic competition stressing the wasteful aspects of competition. Kaldor (1935, p. 34) considered their theory of monopolistic competition a revolutionary doctrine which provides “an excellent theoretical background for the age-old cry of businessmen about ‘the wastes of competition’,” which economists have completely neglected before. The theory of monopolistic competition is undoubtedly one of the key achievements of the 20th century microeconomics. At the same time, these early neoclassical economists described excess capacity as an inefficiency of monopolistic competition. Chamberlin (1952) himself viewed excess capacity as an important phenomenon, although he insisted that it was wrongly identified as the departure from the minimum point on the cost curve. His was an “attempt to combine the two theories of monopoly and of competition into a single one which would come closer to explaining the real world, where, it seemed, the two forces were mingled in various ways and degrees” (Chamberlin, 1952, p. 320). Chamberlin (1952) tried to blend monopoly and competition and viewed the marginal revenue curve as a piece of pure technique unrelated to the central problem. He assumed that monopolistic competition shares the features of monopoly but, at the same time, perceived its demand curve to be flat and has “described the typical curve as highly elastic” (Nichol, 1934, p. 135).

Yet, the neoclassical school maintained that the monopolistically competitive firm is inherently inefficient because in a long-run equilibrium it produces at less than its capacity output and achieves zero economic profit. Monopolistic competition is invariably suboptimal to the perfectly competitive firm whose long-run equilibrium is exactly at the capacity output, i.e., at minimum long-run average costs. Monopolistic competition was compared to perfect competition, an ideal outcome which is unrealistic in the presence of transaction costs and product differentiation in today’s real economy. Chamberlin’s claim that the monopolistically competitive firm behaved monopolistically along its demand curve and, therefore, could set its own price turned into a mantra which was consistently restated by economists. The name adopted by Chamberlin to describe imperfect competition under free entry, although paradoxical, became standard terminology in the field. The welfare implication was clear – too much competition can be bad as it wastes resources. Monopolistic competition was seen as a sinful, less efficient market structure.

The concept of excess capacity was immediately adopted in the literature. Cassels (1937) distinguished between excess capacity and overinvestment. He defined excess capacity as the difference between actual output and potential output, the latter being the economic optimum. There were as well some early opponents of the theory who criticize it on various grounds. The first to attack the concept of excess capacity was Nichol (1934) who stressed the role of marginal buyers with product differentiation. Nichol (1934) demonstrated that if the discontinuous demand curve for each individual product lies entirely below the average cost curve, then no firm has a reason to depart from capacity output. Harrod (1952) maintained that the entrepreneur would choose a plant which avoided excess capacity and would charge a price yielding normal profit. The entrepreneur would plan equipment accordingly, i.e., “on a scale that gives the lowest cost for producing what he can sell at such a price, and, having acquired the equipment, will sell at that price.” (Harrod, 1952, p. 151).

If perfect and monopolistic competition were viewed as close market structures, the possibility for excess capacity is negligible. Hicks (1935) questioned adopting more complicated assumptions if the assumptions of pure competition gave sufficiently close approximations which became the classic Chicago School view on the matter. Machlup (1939) claimed that in the case of monopolistic competition, i.e., competition among many sellers with differentiated products, but without any oligopolistic grouping, the market results related to prices and output would not be much different from those of pure competition. Schumpeter (1939) and Nicols (1947) suggested that perfect and monopolistic competition differed in nothing but product differentiation. According to Nicols (1947, p. 33) the competitive equilibrium and the Chamberlin-Robinson equilibrium are essentially the same except that “consumers distinguish between products of the same general class” where Robinson (1933, p. 89) recognizes that while large numbers are likely to be realized, “the existence of a perfect market is likely to be extremely rare in the real world.” Thus, she recognized that true, real-life competition is monopolistic, not perfect. Nicols (1947) also found that many of the cases treated by Chamberlin and Robinson are oligopolistic or monopolistic situations. Nutter (1955, p. 527) argued that varieties are “pure” substitutes for each other even though they may not be “perfect” substitutes. Thus, although differentiated, products sold in monopolistically competitive markets are essentially the same.

These early attempts could be attacked on various grounds, as imperfectly competitive markets evolved with time. While he admitted that monopolistically competitive markets resemble perfectly competitive ones, Machlup (1939) suspected that purely monopolistic competition is less common, which reduced the practical usefulness of what he called polypoly with differentiated products.[[3]](#footnote-3) This, in his view, makes oligopoly more prevalent than monopolistic competition. It probably took some time for this idea to crystalize since in 1952 Machlup came up with a slightly different interpretation. In his book “The Economics of Sellers’ Competition” Machlup (1952) reiterated that the theory of monopolistic competition is not much different from that of perfect competition in that the demand curves are almost horizontal rather than perfectly horizontal. This was the first hint in the literature at the geometrical significance of the demand curves for the comparative analysis of the two market structures.[[4]](#footnote-4)

Machlup (1952) also did not seem to be at ease with the “wastes of competition.” He called for cost studies in selected industries which would allow moving away from pure speculation about the “wastes of competition” to a meaningful evaluation of the actual situation. Until empirical evidence was provided, he was inclined to take the “wastes” story lightly. That demand curves with monopolistic competition are very flat turned out to be a major attack on the theory of monopolistic competition and its “wastes,” as identified by Robinson (1933) and Chamberlin (1933, 1947). This is because if curves are very flat with slightly differentiated products (and consumers might see the products as identical), then there is little potential for unused capacity and the individual firm will be operating very close to its minimum efficient scale. Professor Machlup’s second contribution was the emphasis on the graphical apparatus. Neoclassical writers seemed to be at ease with the flatness of the demand curve when it came to perfect competition but were uncomfortable discussing the slope of the demand curve for “the large group” case. Product differentiation, and particularly advertising as part of the promotional mix of the firm, were new elements that could not be easily accommodated in existing models. With the introduction of product differentiation economists were no longer inclined to use graphs to depict the point of tangency of the demand and average cost curve in the long-run equilibrium of the monopolistically competitive firm. Chamberlin’s perception that the demand curve of the monopolistically competitive firm was very flat contradicted his own assumption that monopolistic competition shared the features of monopoly.

Machlup’s observations about the geometrical inconsistencies of excess capacity from his book “The Economics of Sellers’ Competition” published in 1952 remained somewhat unnoticed. One possible reason could have been the forcefulness of the proponents of excess capacity who full-heartedly continued to propagate that monopolistic competition firms operate at less than its optimal size. The concept of excess capacity gained a momentum in the academic circles and started to be widely taught in economic courses. Another problem could have been the length and sophistication of the book’s exposition. Machlup also seemed to doubt the steepness of both the demand and the average cost curves. While a flat demand curve for the individual monopolistically competitive firm seems to refute the excess capacity theorem, the relatively flat and extended average cost curve, one of large indivisibilities, seems to reinforce the ’wastes’.

More than a decade later, Baumol did pick up on Machlup’s observations. In his article “Monopolistic Competition and Welfare Economics” Professor Baumol emphasized the validity of Machlup’s findings and restated the geometrical paradox associated with the “excess capacity theorem.” Demand curves are either relatively flat and low with free entry and slight product differentiation or quite steep with significant product differentiation and the marker power it provides to oligopolistic firms. In evaluating Chamberlin’s excess capacity theorem Baumol (1964) stressed that in order for the theorem to be valid, firms in the industry should be sufficiently similar, small and numerous. Thus, the range of relevance of the theorem was restricted severely not only because the assumptions are far from universal, but also because the demand curve of the typical firm is expected to be quite flat and excess capacity correspondingly insignificant.[[5]](#footnote-5)

Baumol emphasized the tautological character of the theorem. If costs are exclusive of rents, then there is no point of tangency and consequently no excess capacity. If, however, rents are incorporated into the average cost and the long-run envelope curve is comprehensive of all costs, then economic profit is not depicted, and the monopolistically competitive firm would be at the point of tangency. Accounting for economic rent raises the demand function to a new level where price and average cost are tangent.[[6]](#footnote-6) Robinson’s statement that demand curves will be tangential to cost curves and that firms will be of less than their optimal size was reduced to a tautology (Kaldor, 1935). Baumol stressed that rents could be divided into those which are socially “allowable” and those which do not increase social product and are, therefore, socially undesirable. He gave the example of a firm which extracted monopoly rent from a scarce resource such as a unique location. Then the scarce resource constituted an advantage to both the firm and the community which is why the rent was an allowable element of the social average cost curve. It was, therefore, justified to incorporate this allowable rent into the total economic costs of the firm. But product differentiation can create a fictitious or illusory quality advantage, for instance, by salesmanship, where there is no real gain to society and, hence, the social average cost curve need not be tangent to the price function, so there might be any excess capacity.

Baumol (1964, p. 48) concluded that where rents are real, excess capacity must also be real, while in the case “where we may escape true excess capacity, we are likely to become tributaries to the artful producer of artificial monopolistic elements.” But monopoly rents understandably accrue to monopoly firms employing unique resources, while monopolistically competitive firms emphasize product differentiation, real or illusory, more than anything else. With free entry and exit, monopolistically competitive firms are less likely to accumulate excessive economic profits which reaffirmed Baumol’s conclusion that excess capacity was hardly present with imaginary monopolistic elements. Baumol seemed to realize that excess capacity came with true market power and not with free, though imperfect, competition.

Baumol (1964) criticized Demsetz (1959) who attacked the excess capacity theorem on account of the multiplicity of variables which enter the profit function under monopolistic competition. Baumol reminded him that there cannot be optimum with a negatively sloping average cost curve. If the firm increases advertising and output, its average cost will rise accordingly which reinforced Chamberlin’s argument that the equilibrium point is characterized by unused opportunities for cost reduction. Baumol’s response to Archibald (1961), who claimed that the average cost curve which is downward sloping at the equilibrium point is an aggregation of production and advertising was similar in character. According to Archibald (1961, p. 19) "when advertising is introduced, tangency is consistent with production at less than or more than minimum cost of production." Baumol observed that Archibald did not try to draw any welfare implications. Baumol then cited Friedman (1963) who did not attribute much significance to the theorem and the optimal scale of operations. What mattered to Friedman was the revaluation of specialized resources by the capital market. This, in Friedman’s view, rendered the tangency condition irrelevant to productive adjustments and the implications of the excess capacity theorem of unexploited economies of scale highly misleading, if not wrong. According to Stigler (1937) the theory of imperfect competition has raised questions which it could not answer satisfactorily until the theory of perfect competition was developed more fully. Therefore, perfect, not monopolistic, competition should be the focus of study of economic theorists. The Chicago School viewed the theory of monopolistic competition as a departure from the strict economic analysis which economics ought to follow. The theory represented a critique of the market system which would lead to calls for government intervention in the product markets. While Stigler opposed the theory because it would render economic analysis more case-oriented, Friedman (1953) questioned it on methodological grounds (Tsoulfidis, 2009).

Baumol and Machlup had the same intuition about excess capacity and “the wastes of competition” probably because they shared a similar past. Both had parents who were small proprietors in competitive industries. Machlup’s father owned a cardboard factory, while Baumol grew up among small retailers. Baumol (1964, p. 50) could easily see how small firms selling heterogeneous products tended to drive each other’s profit down to zero without monopoly rents at an “uneconomically small scale of operation.” He observed that the tangency analysis could provide a meaningful hypothesis to explain this phenomenon but stressed the significance of “an empirical hypothesis about the nature of entry under monopolistic competition, which, if valid, will prevent any substantial monopoly rents and hence any material revaluation of resources by the capital market.” At the same time, the assumption of free entry and exit balances things out. Baumol (1964, p. 51) was aware that demand may be insufficient to allow the individual firm to sell at the minimum efficient scale and at least average cost:

“Some companies will then be forced to leave the field and the demand curves of the remaining firms will shift to the right. This process will continue until each remaining firm can just sell the least cost output and there the equilibrium point will be found.”

Baumol took a global, economy-wide perspective in evaluating the welfare implications of the theory of monopolistic competition. He went on to summarize the two chief reservations on the implications of what he now called “the excess capacity hypothesis.” The first reservation is on industry size and configuration:

“The excess capacity theorem is not a statement about the desirability of the allocation of resources among industries. It does not say that there will be too little produced by an industry (however defined) whose products are differentiated. Rather, the theorem tells us that the organization of the "industry" into firms is apt to be wasteful. It suggests that the same total output if produced by a smaller number of more sizable firms, can be provided at a lower real cost per unit, and hence a smaller total use of society's scarce resources.” (Baumol, 1964, p. 50).

In the case where firms use capacity fully, we are condemned of monopoly rents. If they compete freely bringing monopoly profits all the way down to zero, we are condemned of unused capacity. The theory of monopolistic competition became a justification for the existence of large firms or the introduction of artificial entry barriers. Assuming that the number of firms in the industry is excessively large and far beyond what would be needed for achieving the efficient industry output, one needed to introduce restrictions or find proper ways to handle this excess capacity.[[7]](#footnote-7)

The second key reservation to the implications of the excess capacity theorem referred to product variety and consumer choice. Discussing the social costs of standardization, Baumol (1964, p. 51) observed that if the number of firms in the industry was reduced, the variety of products available to consumers must fall. The resulting saving in resources is then to be considered a net gain depending on the case and “the excess capacity theorem represents a real social cost only if the increased choice which it offers consumers does not provide adequate psychic compensation for the added costs which it imposes.” In other words, product variety, which brings satisfaction to consumers, should be evaluated at the margin. Variety is justified if it brings sufficient happiness to consumers. Standardization and reducing the number of firms deprive consumers of utility which comes with product variety and monopolistic competition. Variety is invariably related to free competition.[[8]](#footnote-8)

Baumol did come up with prescriptions for the elimination of excess capacity. In industries where excess capacity existed it could be eliminated by direct nationalization, a licensing system or price controls. With the help of a licensing system the government could restrict the number of firms to a level where each surviving firm will face a higher demand curve and eventually reach its capacity output. Price ceilings could be used to induce firms to produce at the minimum efficient scale, an idea Baumol borrowed from Sandberg (1963). These measures completely contradicted the perception of deregulation and privatization associated with the contestability theory.

In his later research on contestability Baumol and his team returned to the study of scale economies. They investigated those in the context of the multiproduct firm, natural monopoly, and economies of scope. Baumol et al. (1982) did not consider scale economies and large setup costs a barrier to entry. Rather, it was sunk costs which erect substantial barriers to entry. In the absence of sunk costs, the hit-and-run strategy allows entry even in monopolistic and oligopolistic markets. According to Mosca (2008) Baumol’s group replaced the notion of perfect competition with that of contestability so even in market structures exhibiting monopoly power firms would be forced to behave competitively.

Baumol recapitulated that more has been expected of the newborn and more general theory of monopolistic competition. He attributes the paucity of the welfare results associated with the monopolistically competitive firm to the poor general state of welfare theory. According to Baumol Chamberlin’s book had expanded the economists’ view of market structures. Economists have realized that there was a wide variety of monopolistic elements and characteristic features in different market structures which may have serious social consequences. Restricting the analysis to just the two extreme forms, pure monopoly and pure competition, and ignoring the in-between situations limits the relevance of the welfare conclusions of industrial organization. In his attempt to critically analyze and summarize alternative theories and views of the excess capacity theorem he rediscovered that with free entry and slight product differentiation monopolistically competitive firms were not much different from perfectly competitive ones because their individual demand curves are relatively flat and low given the size of the market. Thus, monopolistic competition was more of a competition, than monopolistic.

Concerned about the state of welfare economics, Baumol wrote his book “Welfare Economics and the Theory of the State.”[[9]](#footnote-9) Based on his PhD dissertation the book came out in its first edition in 1952. He discussed why citizens would voluntarily subject themselves to the coercive power of the state which is not necessarily irrational behavior. According to Baumol some policy measures of welfare theory are examples of a coercive public sector such as taxation, national defense, counter-cyclical measures, and traffic congestion problems. Baumol was also concerned with externalities (Noci, 2019). In the presence of externalities, the perfectly competitive outcome was not a benchmark of optimum production. Externalities rendered market structures such as monopoly more difficult to evaluate and compare to the social optimum. In fact, even the desire for profit and the analysis of profit maximization became contested in the presence of externalities. In the second edition of “Welfare Economics and the Theory of the State” of 1965 Baumol was more concerned with the application of welfare economics to a variety of problems of public policy. He believed that welfare analysis should be more empirically grounded and applied, an idea shared by Machlup (1952). Baumol did turn to empirical studies after that. Baumol and Braunstein (1977) studied scale economies and production complementarity in the case of journal publication. Baumol and Blackman (1980) investigated the environment in the context of the energy sector, while Baumol and Blackman (1983) discussed electronics and libraries in relation to the cost disease. In 2001 Baumol and Wilson (2001) published a comprehensive, three volume collection of the major writings in the field of welfare economics. In “Welfare Economics” the two authors collected seminal research in the field by prominent scholars starting from Vilfredo Pareto and Paul Samuelson all the way to Ronald Coase and Jean Tirole. They included their own writings on externalities, optimality, marginal cost pricing, predation, contestability, equilibrium, and adverse selection.

3. Conclusion

Early authors on the subject of monopolistic competition such as Robinson (1933) and Chamberlin (1933) saw excess capacity as one of the most prominent wastes of competition. They considered the number of firms in the industry excessive while competitive industries were blamed for overinvestment and overproduction. Firms experienced idle capacity. In 1952 Machlup expressed the idea that in monopolistic competition demand is highly elastic which does not allow the individual firm to charge its own price, as claimed by the proponents of the excess capacity theorem. The theory was rediscovered in 1964 by Baumol in his comprehensive welfare analysis of the theory of monopolistic competition. He demonstrated that the monopolistically competitive firm could not deviate substantially from the capacity output in a long-run equilibrium. With very flat demand curves, sellers cannot be price setters. Baumol emphasized that economic rents can be socially allowable and socially insignificant. Where rents are real, as in monopoly and oligopoly, excess capacity must also be real, and costly for society. With free entry and exit, monopolistically competitive firms were unlikely to accumulate rents and excess capacity was not likely to be present. Baumol’s contribution to welfare theory is illustrating the inconsistencies of the “excess capacity theorem” and the “adverse” welfare effects of monopolistic competition. He argued that it was suboptimal to introduce artificial barriers-to-entry to prevent the “wastes” of competition. Baumol suggested that, if present, excess capacity in different industries could be eliminated by direct nationalization, a licensing system which would reduce the number of firms to a level preferable for the authorities, or the introduction of price ceilings which will bring surviving firms closer to the capacity output.

Contestability theory, which Baumol is best known for, leaves many with the impression that Baumol was an advocate of firms with market power, which would lead the industry to the competitive outcome without regulation and government control. In the absence of entry and exit barriers stay-out pricing and other strategic pricing tools would keep new entrants out of the monopolistic industry. We have demonstrated that Baumol was a strong supporter of competition of any kind. He defended monopolistic competition, an industry without any barriers, entry or exit. Based on the analysis of flat demand curves, insignificant or nonexistent excess capacity and product variety, Baumol justified competitive market structures.

References:

Archibald, G. (1961). Chamberlin Versus Chicago. *The Review of Economic Studies*, 29(1), 2-28. Retrieved December 27, 2020, from <http://www.jstor.org.libproxy.aubg.bg/stable/2296178>.

Bailey, E., and W. Baumol. (1984). Deregulation and the Theory of Contestable Markets. *Yale Journal of Regulation* 1: 111–137.

Baumol, W.J. (1952). *Welfare Economics and the Theory of the State*. London: Longmans, Green and Co.

Baumol, W.J. (1964). Monopolistic Competition and Welfare Economics. *American Economic Review*, Vol. 54, No. 3, Papers and Proceedings of the Seventy-sixth Annual Meeting of the American Economic Association, (May, 1964), 44-52.

Baumol, W.J. (1965). *Welfare Economics and the Theory of the State*. Cambridge: Harvard University Press, second edition.

Baumol, W.J. (1982). Contestable Markets: An Uprising in the Theory of Industry Structure. *American Economic Review* 72: 1–15.

Baumol, W.J., and Y. M. Braunstein. (1977). Empirical Study of Scale Economies and Production Complementarity: The Case of Journal Publication. *Journal of Political Economy*, *85*(5), 1037–1048. <http://www.jstor.org/stable/1830344>.

Baumol, W.J. and S.A.B. Blackman. (1980). Unprofitable Energy Is Squandered Energy. *Challenge*, *23*(3), 28–35. <http://www.jstor.org/stable/40719881>.

Baumol, W.J. and S.A.B. Blackman. (1983). Electronics, the Cost Disease, and the Operation of Libraries. *Journal of the American Society for Information Science*, Association for Information Science and Technology, 34(3), 181-191.

Baumol, W.J., J. Panzar, and R. Willig. (1982). *Contestable Markets and the Theory of Market Structure*. New York: Harcourt Brace Jovanovich.

Baumol, W.J., J. Panzar, and R. Willig. (1983). Contestable Markets: An Uprising in the Theory of Industry Structure: Reply. *American Economic Review* 73: 491–496.

Baumol, W.J, J. Panzar, and R. Willig. (1986). On the Theory of Perfectly-Contestable Markets. *New Developments in the Analysis of Market Structure*. ed. J. Stiglitz and F. G. Mathewson. Cambridge: MIT Press, 339–365.

Baumol, W.J., and R. Willig. (1986). Contestability: Developments since the Book. *Oxford Economic Papers*, 38, new series, 9-36. Retrieved December 30, 2020, from <http://www.jstor.org.libproxy.aubg.bg/stable/2663000>.

Baumol, W.J. and C. Wilson. (2001). *Welfare Economics*. International Library of Critical Writings in Economics, Edward Elgar Publishing.

Cassels, J.M. (1937). Excess Capacity and Monopolistic Competition. *The Quarterly Journal of Economics*, 51(3), 426-443. Retrieved December 25, 2020, from <http://www.jstor.org.libproxy.aubg.bg/stable/1884835>.

Chamberlin, E.H. (1947). *The Theory of Monopolistic Competition*, Cambridge: Harvard University Press, 5th ed. [Originally published in 1933.].

Chamberlin, E.H. (1952). “Full Cost” and Monopolistic Competition. *The Economic Journal*, Vol. 62, No. 246 (June, 1952), 318-325.

Demsetz, H. (1959). The Nature of Equilibrium in Monopolistic Competition. *Journal of Political Economy*, Vol. 67, No. 1, 21-30.

Friedman, M. (1953). *The Methodology of Positive Economics*. Chicago: Chicago University Press.

Friedman, M. (1963). More on Archibald Versus Chicago. *The Review of Economic Studies*, 30(1), 65-67. Retrieved December 28, 2020, from <http://www.jstor.org.libproxy.aubg.bg/stable/2296033>.

Harrod, R.F. (1952). Theory of Imperfect Competition Revised, in *Economic Essays*.

Hicks, J.H. (1935). Annual Survey of Economic Theory: The Theory of Monopoly. *Econometrica*, Vol. 3, p. 12.

Kaldor, N. (1935). Market Imperfection and Excess Capacity. *Economica*. New Series. Vol. 2, No. 5, 33-50.

Machlup, F. (1939). Evaluation of the Practical Significance of the Theory of Monopolistic Competition. *The American Economic Review,* *29*(2), 227-236. Retrieved December 24, 2020, from <http://www.jstor.org/stable/1803621>.

Machlup, F. (1952). *The Economics of Sellers’ Competition*. Johns Hopkins.

Machlup, F. (1974). Friedrich Von Hayek’s Contribution to Economics. *The Swedish Journal of Economics*, 76(4), 498–531. <https://doi.org/10.2307/3439255>.

Mosca, M. (2008). On the Origins of the Concept of Natural Monopoly: Economies of Scale and Competition. *The European Journal of the History of Economic Thought*, 15(2), 317-353, DOI: 10.1080/09672560802037623.

Nichol, A.J. (1934). The Influence of Marginal Buyers on Monopolistic Competition. *The Quarterly Journal of Economics*, Vol. 49, No. 1 (Nov.,1934), 121-137.

Nicols, A. (1947). The Rehabilitation of Pure Competition. *The Quarterly Journal of Economics*, Vol. 62, No. 1, (Nov. 1947), 31-63.

Noci, A. (2019). William J. Baumol and the New Welfare Economics Debate: 1940 – 1970. Paper presented at the *European Society for the History of Economic Thought Conference “Money, Banks, and Finance in Economic Thought,”* University of Lille, May 23-25, 2019.

Nutter, G.W. (1955). The Plateau Demand Curve and Utility Theory. *Journal of Political Economy*, Vol. 63, No. 6 (Dec., 1955), 525-528.

Robinson, J.V. (1933). *The Economics of Imperfect Competition*. London: Macmillan, 2nd edition, 1969.

Sandberg, L. (1963). A Note on the Effect of Price Controls on the Theory of Monopolistic Competition. *The American Economist*, 7(1), 48-53. Retrieved December 28, 2020, from <http://www.jstor.org.libproxy.aubg.bg/stable/25602591>.

Schumpeter, J.A. (1939). *Business Cycles*, New York: McGraw-Hill, Vol. I.

Stigler, G. (1937). A Generalization of the Theory of Imperfect Competition. *American Journal of Agricultural Economics*, Agricultural and Applied Economics Association, 19(3), 707–717.

Tsoulfidis, L. (2009). The Rise and Fall of Monopolistic Competition Revolution. *International Review of Economics*, 56(1), 29-45, DOI:10.1007/s12232-008-0051-8.

1. Baumol et al. more than once cited the airline industry as an example of contestable markets where investments in aircraft do not incur any sunk costs because they constitute “capital on wings” which could easily be recuperated by entrants upon exit of the airline business. Later, the authors considered transportation by trucks, barges, and buses even more highly contestable than passenger aviation (Bailey and Baumol, 1984, p. 24). [↑](#footnote-ref-1)
2. Baumol and Willig (1986) claimed that the contestability theory was never intended to remove all regulation or antitrust activity because it is a source of economic inefficiency, or to imply that the unrestrained market solves all economic problems: “… it is simply incorrect to associate our writings on contestability with an all-pervasive laissez-faire position on the role of regulation and antitrust.” (Baumol and Willig, 1986, p. 10). [↑](#footnote-ref-2)
3. In his 1939 article “Evaluation of the Practical Significance of the Theory of Monopolistic Competition” Machlup (1939, p. 231) stated, “I have another suspicion which might reduce the practical usefulness of this pattern of monopolistic competition, that is, of the pattern of polypoly with differentiated products. I suspect that cases of this type are less frequent in reality than has been assumed, because where products are differentiated there are usually smaller groups of closer rivalry within the market of many sellers. This brings an element of oligopoly into almost every market; the market of supposedly very many competitors is broken up into many small and partly overlapping groups of rivals whose position is clearly oligopolistic.” [↑](#footnote-ref-3)
4. Machlup (1952, p. 316, italics in the original) wrote, “In this last interpretation the results of combining the theories of *imperfect* polypoly and perfect pliopoly are not much different from those reached by the synthesis of *perfect* polypoly and perfect pliopoly. In the geometric representation, the difference is merely that the demand curves are *almost* horizontal rather than *perfectly* horizontal. The thesis of the combined theories, that the output produced will fall short of the capacity output, and the size of the firm will be below the optimum size, will hold true, but it will not be quantitatively important. The deviation from the optimum would amount to much only if the demands were not so highly elastic as we have concluded they must be for the tangency rule to apply at all, according to this, the most reasonable interpretation. With the highly elastic demands the deviations will be negligible.” [↑](#footnote-ref-4)
5. Machlup (1952, p. 314) described the situation very subtly: “Thus, while the products are different enough to set limits to the salability of anyone of them, that is to say, to prevent the demand for anyone product from being perfectly elastic, they are still similar enough to make the demand for anyone product so highly elastic that the seller has only little choice concerning the price he can charge. The more elastic the particular demand curves, the smaller will be the range of possible prices among which the seller can choose in any given situation. If the demand curves are shifted under the pressure of an increased supply of similar products and if they maintain their high elasticity (and I dare say there are good reasons for assuming the elasticities to be still further increased by the newcomers' competition) the adjustments of sales expectations will predominantly be adjustments of price.” [↑](#footnote-ref-5)
6. Baumol (1964, p. 47) wrote, “The profits of the firm under monopolistic competition must be imputed to one or more of its inputs, and if this rent is included in cost, the company's net profit must necessarily be zero. The tangency result and the excess capacity theorem therefore follow inescapably. It is worth noting that Mrs. Robinson's version of the theorem is of this variety, which, as Machlup has shown, is not entirely without welfare implications, despite its tautological character.” [↑](#footnote-ref-6)
7. In providing examples of excess capacity Cassels (1937) mentions mills standing idle, factories working shortened shifts, buildings unoccupied, ships laid up in dock, etc. Many of these examples referred to “the small group case”, rather than “the large-group case” and often represented short-run situations. It is also worth noting that empty factories and unused machinery are often not a social cost, but a private loss to entrepreneurs who overinvested. [↑](#footnote-ref-7)
8. Some scholars prematurely concluded that rational consumers would immediately opt for the cheaper, standard product. As Machlup (1952, p. 323) observed: “There are those who take it for 'granted that the public would decide in favor of the cheaper products. They are apt to consider the differentiation as a sheer waste that ought to be prevented by deliberate action of the government. Since newcomers' competition causes the waste of excess capacity, restrictions of entry appear to them as a logical remedy… Restrictions of output of existing firms are to be combatted by introducing restrictions of entry of new firms. This, in my opinion, is very bad policy...” [↑](#footnote-ref-8)
9. For a thorough analysis of Baumol’s contribution to welfare economics and the theory of externalities, see Noci (2019). [↑](#footnote-ref-9)