accomplished by different firms merely assigning different brand names to identical substances and then advertising such brand names intensively; and (4) the presence of a very high degree of concentration in the production of most of the important ethical drugs, in the absence of any apparent economies of large-scale production.

An investigation of the production and marketing of ethical drugs compels the conclusion that the industry may easily be considered delinquent in regard to its impact on the allocation of economic resources. This paper will attempt to demonstrate that such misallocation of resources has been made possible by the existence and abuse of the patent privilege for ethical drug products and processes and by a number of measures which the industry has taken to foster and exploit the remarkable degree of imperfection of market information which the structure of the ethical drugs market permits. Such measures can, however, be understood adequately only in the context of actual drug marketing practices, which will be described below.

I. THE ECONOMIC FRAMEWORK OF THE ETHICAL DRUGS INDUSTRY

Broadly speaking, the economic framework within which the ethical drugs industry operates is conditioned by the nature of demand, the structure of costs, the existence of the patent privilege for drug products and processes, the eligibility of branded drug products for protection under the trademark laws, and the existence of the Food and Drug Administration's drug product inspection facilities as an external means of insuring quality in drug products.

The nature of demand is such as to distinguish the prescription drugs market from virtually all other markets.[3] Sole purchasing authority, as well as the necessary initiative, lies with the prescribing physician, who orders the specific drug for which the patient must pay. Chemically identical drugs may be selling under different names at widely varying prices, but the physician has no direct motivation to prescribe the lowest-priced brand or even to become aware of prices at all. (Indirectly he may reason that his fee for services may more readily be capable of collection if drug prices are low, but this would seem to be a distinctly secondary consideration.) Ideally, the total potential market for a drug or for a group of related drugs consists of the total need for medication (either for cure or for the relief of symptoms) on the part of all individuals afflicted by the various disease entities or clinical syndromes which are capable of treatment by the drug or group of drugs. Economically, total effective demand at given drug prices may fall short of total physical need in the case of individuals with very low incomes and no access to welfare case status, and it may exceed total physical need to the extent that individuals not suffering from those disorders for which the drugs are of use may nevertheless be treated with them. For any given drug, therefore, the actual relevant market is comprised of the total effective demand for medication on the part of all individuals who can be induced to consult physicians and who are afflicted by those disorders for which physicians may be inclined or persuaded to prescribe the particular drug.

Either for single drugs, or for a group of related drugs as a whole, the demand curve is likely to be extremely inelastic. It has been reported that individuals with severe inflammatory diseases and low incomes sometimes do without food in order to buy drugs.[4] Shifts in the demand curves for individual products are brought about by direct-mail advertising, medical journal advertising, and by the insistence of itinerant sailesmen or "detailmen," employed by the major drug firms. Advertising cannot manipulate the total incidence of diseases (although to a limited extent, news articles in newspapers and magazines may mention that a given drug can treat a given condition, and thus make more people who suffer, or imagine that they suffer, from a given condition, aware that drug treatment is available); but it can shift the existing effective demand from one product to another, and such shifts can perhaps be visualized as parallel rightward or left-

[3] Exception might be taken to this statement by college professors, who may be said to "prescribe" textbooks for their students. Here, however, the student has alternatives not open to drug buyers, such as the existence of a second-hand textbook market and the possibility of textbook sharing by two or more students.

[4] Testimony of E. D. Bransome of the Arthritis and Rheumatism Foundation. Hearings on Administered Prices before the Subcommittee on Antitrust and Monopoly of the Senate Comm. on the Judiciary, 86th Cong., 1st Sess., pt. 14, at 7992–93 (1960). (Hereinafter cited as "Hearings on Administered Prices.")

ward movements of nearly vertical demand curves. The absence of price-consciousness on the part of the physician, the inability of the patient to purchase any but the specified drug, and the marked inelasticity of the individual patient's demand curve, all are conducive to the possibility of charging a price which is extremely high relative to marginal cost.

The cost structure of the industry is such that there seem to be few, if any, notable economies of scale in the production process. In this regard, a sharp distinction must be drawn between the production of "heavy" or "bulk" chemicals which are commonly produced in enormous quantities by continuous-flow processes (sulfuric acid would be a good example) and the production of "fine" chemicals, such as pharmaceuticals. Batch methods, which allow far less scope for economies of scale, are the rule, especially where fermentation is the key phase in the production process, as is the case in the production of antibiotics (except chloramphenicol) and synthetic corticosteroid hormones.[5] These two categories of ethical drugs are the largest in terms of total sales volume. Fermentation vessels are "amazingly standard." [6] Large antibiotics makers usually employ 10 to 15 such vessels.[7] Changes in output levels for such products may be accomplished by the employment of a differing number of individually identical fermentation vats, a circumstance conducive to constant returns to scale. A second reason for the absence of important economies of large-scale production is the frequent absence of truly large-scale production. The physical volume of output of the active ingredients in drugs is typically very small, total national output being perhaps in the neighborhood of a ton pear year.[8] Actual cost data are kept secret by the drug firms. The inconclusiveness of some of the economic aspects of the 1959–1960 Senate hearings on the ethical drugs industry is due to the "trade secret" status afforded certain crucial cost data and the consequent failure to insist upon the publication of actual production costs for at least some of the higher-priced patented drugs. To the economist, this is a most unfortunate omission, for such data are otherwise entirely unobtainable. On the basis of the costs submitted to the Senate Subcommittee (but not published in the record), the *Report* comments upon the absence of economies of scale and the resulting "extremely small size of plant required for economical production." [9] It is certainly true that very small firms can purchase the active ingredient for a given drug in bulk (the so-called "bulk powder") from the large firms and tablet and sell it at prices greatly below those which the large firms see fit to charge; there is also considerable evidence that very small firms can actually manufacture their own fine chemicals as bulk powder and sell them to small firms in bulk at prices equal to or lower than those charged by the major firms.[10] Similarly, on the basis of antibiotics production costs revealed to the Senate Subcommittee, but not published in the record, Senator Kefauver said at one point in response to Eli Lilly & Company's request for confidentiality as to its production costs, "I will tell you frankly you all have about the same cost figures. . . . For a difference of one or two cents . . . I do not see any sense in all the secrecy. . . . Your costs, Bristol's costs, Upjohn's costs, are all within a very, very narrow range." [11] This tends to support the contention that there are no important economies of scale in production, inasmuch as the level of output of Bristol's product was almost twice as great as Lilly's. Granted that the ethical drug firms can purchase bulk chemicals from heavy chemicals firms which do themselves produce under conditions of economies of scale, within ethical drugs proper, there seems to be no evidence of anything but roughly constant returns to scale in the production process. The most important economies of large size seem to lie in the area of large-scale selling and advertising, to be discussed below. In the absence of important economies

---

[5] Cortisone Quest: The Right Process Bug. Chem. Week, Aug. 25, 1952, as quoted in Hearings on Administered Prices, pt. 14, at 8291.

[6] Gaden, Fermentation, Chem. Engr., April 1956, p. 159.

[7] FTC, Economic Report on Antibiotics Manufacture 118 (1958).

[8] Production in 1958 of the most important synthetic corticosteroid hormones (the second largest category of ethical drugs by sales volume) was, according to the U.S. Tariff Commission, as follows: cortisone, 2.15 tons; hydrocortisone, 3.21 tons; prednisone, .63 tons; prednisolone, 1.4 tons; methylprednisolone, .5 tons. Hearings on Administered Prices, pt. 14, at 8285–86.

[9] Subcomm. Report 4.

[10] According to Dr. Phillip Berke, president, Formet Laboratories, a small producer of drug bulk powder, with estimated sales of about $500,000 for all products, could in 1959 sell bulk prednisone at prices competitive with those charged by Merck and Pfizer, firms with over $200,000,000 in sales of all products. Hearings on Administered Prices, pt. 14, at 8056.

[11] *Id.*, pt. 24, at 14125–26.

of scale in production, however, it might be conceivable that the ethical drugs industry could approximate to the condition of pure competition.

Patent protection is available in the drug field for process improvements and for the development of new products. Although a product patent cannot be obtained without an accompanying process patent, this is no obstacle to product patents. "Patent medicines" are by no means peculiar to the proprietary drugs field. Unrestricted patent privileges on drug products are, however, virtually peculiar to the United States. Most countries do not award product patents for drugs. Only 28 of the 77 countries for which the Senate Subcommittee could obtain patent law data granted pharmaceutical product patents. Twenty-five of these countries have provisions for compulsory licensing. Only Panama, Belgium, and the United States are without any such limitations.[12] In the United States, the definition of a patentable drug has been broadened in recent years. Prior to 1946, naturally occurring substances could not be patented; product patents on cortisone and hydrocortisone, for example, were impossible to obtain. In 1946 a patent was obtained on the antibiotic streptomycin, on the grounds that, although it was a product of nature, it was transitory in nature, had never been isolated, and its therapeutic use was unknown.[13] In 1955 a patent was issued for tetracycline, even though this compound was jointly produced in the yields from the process for making chlortetracycline, a previously patented drug, and even though tetracycline had previously been marketed by another firm (which did not receive a patent).[14] Patents, therefore, are becoming increasingly easy to obtain in the ethical drugs industry. Trademarks, in the drug industry as elsewhere, may be employed to extend a privileged market position for a larger period of time than the term of the relevant patent. A good example is the case of the barbiturate phenobarbital, which the patent holder sold as "Luminal", at $9.80 per ounce. After the patents expired in the 1930's, competitors produced phenobarbital for sale at $.85 to $.90 per ounce, but for years "Luminal" continued to sell at $9.80.[15]

The Food and Drug Administration has numerous responsibilities with regard to ethical drugs, chief among which are the control of the composition and quality of ethical drugs, the investigation of the safety of new drugs, the certification of antibiotics and insulin, and the prevention of the illegal distribution of prescription drugs. The laws are very complex. For the purposes of this paper, it may be sufficient to note that the purely economic impact of such regulation is, first, to supplement the quality control of individual firms through the authority of the Food and Drug Administration to inspect samples of ethical drugs (and its more limited authority to inspect production facilities) and, second, to prevent the marketing of dangerously lethal new drugs.[16] Violative samples of drugs may be simply confiscated and destroyed, or legal proceedings may be instituted, and civil and criminal penalties may be imposed by the court, including fines large enough to put a small firm out of business. The certification of new drugs, if done with proper regard for the standards necessary to insure drug safety, can be a time-consuming process. Economically, the consequences of regulation are therefore to provide an additional guarantee to the consumer that all drugs on the market are of acceptable quality and to delay the introduction of new drugs in the interests of safety.

II. DRUG INDUSTRY RESPONSE TO THE SPECIFIC ECONOMIC FRAMEWORK

The ethical drug industry's response to the given economic framework may be interpreted as a successful strategy to exploit the monopolistic potentialities inherent in trademarks, the patent privilege, and the great inelasticity of market, in such a way as to overcome the potential for competition implicit in the absence of economies of scale and, to a lesser extent, in the presence of governmental inspection to insure the safety of all drugs placed on the market.

---

[12] The vast majority of European drug discoveries have been made in countries where no patent protection was available. Germany, Switzerland, Sweden, Austria, the Netherlands, and Italy are without product patents. France had none until 1959. Great Britain allows product patents on new drugs, but not on combinations of known drugs, and all such patents are subject to compulsory licensing at reasonable royalties. Subcomm. Report 105–6.

[13] *Id.*, at 141.

[14] *Id.* at 146.

[15] Hearings on Administered Prices, pt. 18, at 10433.

[16] Testimony of Food and Drug Administration Commissioner G. P. Larrick, *id.*, pt. 22, at 12110–36.

The most prominent monopoly element in the market, the patent privilege, may, as is well known, be employed in many ways to limit competition beyond what is explicit in the specific grant of an individual product or process patent. In the chemicals industries the necessary expedients are of easier access than is generally the case. Original patent grants may be extended indefinitely by a judiciously timed succession of improvement patents.[17] Improvement patents in chemicals are particularly readily devised, especially where there are process patents on a continuous-flow production method. On the other hand, it is relatively easier to "patent around" existing patents by engineering small chemical changes in product composition. Whenever the product or process developed as a result of patenting-around efforts is more efficient than that developed by the original patent holder, a situation results which requires cross-licensing.[18] The negotiating of cross-licensing agreements has as an attendant hazard the introduction of the possibility of such meetings of the minds as may be incident to the making of agreements involving the mutual compromise of patent monopoly positions. Patents in the drug industry may bring about few absolute monopolies (few patented drugs are uniquely efficacious, even for a single disorder) but may provide the occasion, and the economic basis, for the development of a greater sense of community of interest in matters of price and production. So much is speculation. However, the terms of the actual patent licensing and cross-licensing agreements obtained by the Subcommittee, together with the evidence of price uniformity (not only between different brands of the same drug but also among different drugs in the same field of therapy), lend some support to such speculation.[19]

It is of course claimed by the drug makers that patents promote competition. Several relatively large drug firms have asserted that patents are a small company's best friend.[20] By this it is meant that the company in question gained monopoly power in some market and was enabled to out-grow its small, competitive beginnings.[21] This is consistent with the theory implicit in much current anti-monopoly legislation: that the remedy for the existence of monopoly power on the part of the large firms in a given market is to confer a similar degree of monopoly power on the small firms.

Really small firms, on the other hand, have testified that a patent is merely an invitation to costly litigation and the only advantage a really small firm can take of a useful patent is to sell it to a larger rival, usually for a lump sum paid-up royalty.[22] Even if a small firm were willing to face the hazards of litigation, it would lack the funds necessary to introduce even a hypothetical new product of superlative efficacy because of the needed outlay on sales promotion.[23]

---

[17] Insulin is an excellent case in point. See subcomm. Report 141.

[18] This is often the case where the production of a final product requires as inputs a number of intermediate products, some of which may be capable of chemical synthesis as well as of organic cultivation or fermentation. An important example is the case of the broad-spectrum antibiotic, tetracycline, independently discovered by several drug firms at about the same time. Bristol Laboratories used a method of direct fermentation to produce tetracycline. Lederle had patented an earlier antibiotic, chlortetracycline ("Aureomycin") and had never licensed any competitor. Lederle found that by dechlorinating chlortetracycline, the resulting compound, tetracycline, had unusual antibiotic properties. The chemical dechlorination process was more efficient than direct organic fermentation, but Bristol apparently never applied to Lederle for supplies of chlortetracycline, presumably in the interest of secrecy, or in anticipation of refusal. Pfizer had done tetracycline research, and had filed a patent application, but found that Lederle had also done so. Pfizer convinced Lederle that Pfizer's claims were superior; Lederle, however, had patents on chlortetracycline, which Pfizer needed as an input for tetracycline. A mutual agreement was made, insuring both Pfizer and Lederle of a market for tetracycline, regardless of the award of the patent. Hearings on Administered Prices, pt. 24, at 13848. 13697–701.

[19] The currently pending Federal Trade Commission complaint, FTC v. Cyanamid, concerns such a point in connection with the tetracycline patent negotiations mentioned just above. The FTC charged the respondents with a conspiracy in connection with the awarding of the patent to Pfizer, and with a conspiracy in restraint of competition, operating through the subsequent patent licensing arrangements. Subcomm. Report 145–47.

[20] Testimony of Dr. P. I. Bowman, president of Bristol Laboratories. Bristol failed to get the tetracycline patent, but the mere award of a license (in 1956) seems to have been lucrative. Bristol's profits after taxes as a per cent of net worth averaged 2.9 per cent for the four years prior to licensing, but 10.3 per cent for the four years following. Hearings on Administered Prices, pt. 24, at 13835, 13858.

[21] Testimony of Henry Hoyt, president of Carter Products. In the five years prior to the marketing of its patented tranquilizer, meprobamate ("Miltown"), Carter had earned 16.8 per cent after taxes on net worth. In the next five years (1955–1959) the rate of profit after taxes averaged 41.0 per cent. *Id.*, pt. 16, at 9113, 9170.

[22] Testimony of Seymour Blackman, executive secretary, Premo Pharmaceutical Products. *id.*, pt. 14, at 8254.

[23] Testimony of Myron Pantzer, vice-president of Panray Corporation, *id.*, pt. 16, at 9372.

And the most certain evidence of the lack of that sort of competition which is beneficial to the consumer is the rigidity of the prices of patented antibiotics, compared with the steady price decline displayed by the few unpatented antibiotics, such as the earlier penicillins.[24] The extent to which patents are used to limit the number of producers of a given drug can briefly be illustrated by comparing the number of producers for 42 drugs, in the categories of hormones, diabetic drugs, tranquilizers, sulfa drums, and antibiotics. Of these 42 drugs, which account for over half of all ethical drug sales four were unpatented, while 38 were produced under patent or patent-application licensing agreements. Of the unpatented drugs, two were made by three firms, and two by seven firms. Of the patented drugs, 24 were made by a single producer, 8 were made by two firms, 4 by three firms, and 2 by five firms.[25] Twenty-four of 42 drugs or 57.1 per cent, were therefore produced by a single patent holder. The concentration of sales is not as great as the concentration of production, but 16 of the 24 patent holders who license no others to produce, also license no others to sell, and therefore market their products as patent monopolists. One firm which licenses another to produce for it licenses no other sellers, and thus constitutes a seventeenth patent monopolist. Of the remaining eight products for which there is a production monopoly, two producers have licensed only one other seller, three producers have licensed two other sellers, and the two other three producers have licensed six, seven, and eight other sellers.[26] Even in an industry without economies of scale, patents can act as a barrier to entry and enable monopoly positions to be attained.

An inelastic industry demand curve does not in itself insure monopoly power on the part of the seller, if there is workable competition among sellers. In the absence of economies of scale, small firms could compete with large firms on a price basis. To minimize the effect of such competition, the ethical drug industry has taken measures to influence the normal means of dissemination of market information so as to prevent the physician from becoming effectively aware that there are lower priced sellers in the market, to prevent the recognition of lower priced alternatives to higher priced drugs wherever the awareness may exist that there are lower priced drugs, and to convince the physician that all lower priced drugs are of dangerously poor quality. These efforts to maximize the degree of imperfection of market information have been extremely successful, in large measure because of the nature of the market in which drugs are bought and sold. It is essential at this point to examine that market in some detail.

The direct buyers in this market consisted, in 1959, of some 2,800 drug wholesalers, 53,000 retail pharmacies with 110,000 pharmacists, 9,000 public and private general hospitals, and 4,000 veterinary hospitals.[27] The individual patient is the final consumer, but the physician acts as his purchasing agent when writing the prescription. The physician may write the prescription in terms of the generic name of the compound, or by use of a particular firm's brand name. The retail pharmacist may fill generically written prescriptions by using any firm's brand of the drug, but any prescription written in terms of brand names must be filled with the brand called for, unless the pharmacist obtains specific permission from the physician to make a "substitution" of another brand. Unless the pharmacist obtains permission in each instance, he is subject to punitive action by the state pharmacy boards in 44 states under the "substitution" doctrine.[28] It must be admitted, however, that the courts are not always willing to concur in this doctrine. In *Michigan State Board of Pharmacy v. E. L. Casden* the state court dismissed the contention that any meaningful "substitution" had taken place

---

[24] For the ten-year period 1951–1960 the bulk price of penicillin dropped from $2.50 to $0.21 for ten million units; the bulk price of streptomycin dropped from $3.24 to $0.36 for ten grams; the prices of the patented antibiotics chlortetracycline ("Auteomycin"), oxytetracycline ("Terramycin"), chloramphenicol ("Chloromycetin") and tetracycline (introduced in 1955 and sold as "Polycycline," "Achromycin," "Steclin," "Tetracyn," and "Panmycin") remained constant at $5.10 per bottle of sixteen bbilligram capsules, until August 1960 (one month before the scheduled hearings on antibiotics) when they were reduced by 15 per cent. *Id.* pt. 24, at 13664. It should be observed that all of these antibiotics are produced by the same general process of fermentation, with the exception of chloramphenicol, which is produced by a cheaper synthetic process.

[25] Subcomm. Report 67.

[26] *Id.* at 67–68.

[27] Testimony of Dr. Austin Smith, president of the Pharmaceutical Manufacturers Association, Hearings on Administered Prices, pt. 19, at 10728.

[28] Letter from H. S. McNeil, president of McNeil Laboratories, *id.*, pt. 21, at 11713.

when the identical chemical substance was sold under two different brand names.[29]

Most hospitals operate under the formulary system, which authorizes the hospital pharmacist to fill prescriptions with any generic equivalent of the prescribed drug, regardless of the brand name written by the physician. Such formularies operate under agreements involving the written consent of all resident physicians. In practice it also applies to visiting private physicians treating hospital patients. The drug trade associations have attempted to challenge this practice, but without marked success.[30] The lack of success here stands in marked contrast to the effectiveness of the campaign of the National Pharmaceutical Council in influencing state boards of pharmacy to adopt "substitution" laws. "Substitution" used to refer to the dispensing of a different chemical substance than the prescription required. The National Pharmaceutical Council undertook to influence state boards of pharmacy to adopt antisubstitution laws which defined substitution in terms of brand names. In 1953 only four states had any sort of antisubstitution laws, but by early 1959 the efforts of the National Pharmaceutical Council had increased the number to 44.[31]

Hospitals and dispensaries in the Armed Forces operate under the Military Medical Supply Agency, which purchases on the basis of competitive bids. The bids received often fall well below the wholesale price of the drug in some lines. The Veterans' Administration customarily enters into negotiations with individual suppliers, although it has also solicited competitive bids, particularly after the Military Medical Supply Agency had demonstrated how this could lower prices.

The market consists, in addition, of some 170,000 practicing physicians with the M.D. degree (allopathic physicians), 12,000 osteopathic physicians, and 16,000 veterinarians, all of whom act as purchasing agents for the drug consumer.[32] A vast selling effort is directed at these purchasing agents, primarily at the allopathic physicians, to create brand preferences which will result in the writing of brand-name prescriptions. The types of selling effort include (1) advertisements in over 300 medical journals, (2) direct mail advertisements, (3) the employment of chemists and pharmacists as highly trained itinerant salesmen, and (4) the sponsoring of medical conventions and meetings, in connection with the exhibition of wares. For the 22 largest firms in the ethical drugs industry, selling expenses were in 1958 the second largest component of costs, averaging 24.8 per cent of sales, as compared with a cost of goods sold of 32.1 per cent of sales. Average selling expenses were therefore 77.3 per cent of the cost of goods sold. Sixteen companies had a ratio of more than 75 per cent, nine had a ratio of more than 100 per cent, and four had a ratio of more than 150 percent. Total selling outlays amounted to about $3,200 per physician per year.[33]

The result of this volume of selling effort on the structure of the market has been to eliminate price competition except in the hospital and government bidding markets, and to substitute product differentiation and brand preference.[34] Any physician who wishes to put himself to the trouble of looking up the generic name of a compound advertised by brand name can write any prescription generically. The retail pharmacist may then fill the prescription with any available brand. For patented drugs there is little likelihood that there will be any great variation in price, although there are a few such instances.[35] For unpatented drugs, however, the unadvertised, generically named products of small firms may sell for 10 per cent or less of the price of the large firms.

---

[29] Text of Wayne County, Michigan circuit court opinion no. 301799, *id.*, pt. 21, at 11761–62.

[30] Memorandum 18 of the National Pharmaceutical Council, 1957, *id.*, pt. 21, at 11835–36.

[31] Subcomm. Report 236.

[32] Testimony of Dr. Austin Smith, Hearings on Administered Prices, pt. 19, at 10728.

[33] Data submitted by each firm to the Subcommittee. Subcomm. Report 31.

[34] The extent to which advertising, salesmanship, and the bestowal of free samples is regarded by drug makers as capable of creating lasting brand preference is indicated in an article by the vicee-president of Abbott Laboratories, where it is stated: "Practically the only ethical drug products whose sales will not be benefited greatly by sampling [the distribution of free samples] are the rare medicaments distinguished by advantages so great that in simple justice to the patient their actual use is called for." Downs, Advertising, in Drug Research and Development 485 (Smith & Herrick ed. 1948).

[35] Reserpine, a tranquilizer patented by CIBA, is sold under its brand name, "Serpasil," at a wholesale price of $4.50 per hundred .25 milligram tablets. CIBA has licensed several other firms, some of whom make sales of the bulk powder to small companies who in turn can undersell the large firms' advertised brands. The 1959 Drug Topics' Red Book listed 41 firms selling this product. Two firms were selling at over $4.00 per bottle (CIBA's $4.50 was the highest price), two were selling from $3.00 to $4.00 per bottle, two were selling from $2.00 to $3.00 per bottle, eleven were selling from $1.00 to $2.00 per bottle, and 24 were selling at less than $1.00 per bottle. The lowest price, Winsale's $0.45, was exactly 10 per cent of CIBA's price. Hearings on Administered Prices, pt. 18, at 10595.

It is conceivable that some degree of price competition might gain a foothold in the prescription drug market. This is not likely to occur, however, even where generic name products are sold in competition with brand name products, for several reasons. First, the physician is probably only infrequently very price-conscious.[36] It has even been argued that the physician who prescribes the more costly drugs will be accorded the greater prestige. Second, drug advertisements virtually never mention prices, and the various physicians' reference manuals never give competitive prices.[37] It is, in fact, apparently illegal for a generic name product price catalog to compare its prices with those of identical products sold under brand names, or even to mention any of the brand name equivalents of a generic name product.[38] Third, the pharmacist has no economic interest in selling the cheapest brand of drug, since the absolute magnitude of his customary markup of 66⅔ per cent over invoice cost will thereby be minimized. Indeed, it may even happen that the pharmacist will buy generic name drugs and sell them at brand name prices, if the product itself is not clearly identified as to manufacturer, and if the policing policies in use among druggists are not adequate. Fourth, and perhaps most important, the large firms who do the advertising, support the trade associations, and hire the detailmen, make every effort to disparage the products of their smaller competitors who sell at lower prices.[39] There is probably no other industry in existence where the disparagement of the quality of lower priced products can so completely substitute for actual price competition.[40]

Physicians are the chief targets for such "educational" disparagement efforts, but the pharmacists and the general public are also exposed to it whenever opportunity allows. Several representatives from the large drug firms, and from the Pharmaceutical Manufacturers Association, showed no reluctance to use the Hearings as a forum for the disparagement of drugs selling at low prices.

The intent behind these various efforts is clearly to deprive the physician of access to all the information necessary to allow him to function as a competent purchasing agent for his patient, and to influence favorably his attitude toward brand names. To the extent that the physician, who does not pay, can be considered to have a demand curve for the drugs he prescribes, the efforts of the drug makers can be interpreted as attempts either to give the physician's demand curve an upward slope within the relevant region of drug prices, or to reduce his demand to zero for prices lower than those charged by the major firms. The demand curve of the patient is perhaps nearly vertical up to prohibitively high prices if he trusts the judgment of his physician; at any rate, it will be very inelastic. The patient has no alternative but to purchase the medication at the specified price, or go without, and thus forego the benefit of the medical advice he is paying for.

No one wishes to be exposed to the hazards of inferior drugs, regardless of their price. Is there reason to suspect that lower priced drugs are of poorer quality than higher priced drugs? Actually, there is very little to be gained by cutting corners in production and quality control, and much to be lost if such a practice is detected by Food and Drug Administration inspectors. The reasons why no quality differences should be expected to exist between the drugs of large and small sellers may be briefly listed. (1) The same standards of identity, potency, and purity are enforced with regard to all makers for drugs listed in the *United States Pharmacopoeia* and the *National Formulary*. The Food and Drug Administration is empowered to inspect any drug produced by any firm at any time, and to impose criminal penalties upon violators. (2) The Food and Drug Administration has concentrated its drug regulation efforts on the smaller firms. From January, 1950, until June, 1960, 8,376 samples of drugs produced by the 22 major firms were examined, and 8,621 samples from some 1,200 smaller firms were examined. The larger firms produced 87 per cent of all ethical drugs,

[36] Mr. Blackman of Premo testified that it once took him two and one-half hours of argument to convince one physician of his acquaintance that the physician was, in fact, acting as a purchasing agent for his patient. *Id.*, pt. 14, at 8223.

[37] Detailmen are also instructed to ignore prices. One doctor persuaded several of his colleagues to interrogate detailmen on drug prices and costs. If pressed, they would generally quote prices, but no one ever mentioned a cost figure. *Id.*, pt. 18, at 10455.

[38] *Id.*, pt. 14, at 8212–13.

[39] An almost prototypical utterance of this genre is the characterization of the lower priced drug seller as providing "the unfair competition of the unscrupulous one who manufactures under unsanitary conditions, cheapens his product with low quality ingredients, and falsely labels his product as being better than it is. . . ." Statement of Dr. Austin Smith, president of the Pharmaceutical Manufacturers Association, *id.*, pt. 19, at 10738.

[40] See the testimony of Dr. Solomon Garb, *id.*, pt. 18, at 10476–77.

the smaller firms only 13 per cent.[41] The likelihood that a specific drug preparation made by any small firm will be inspected is therefore about 6.9 times as great as in the case of any large firm. It must be conceded, however, that during this period the average large firm had 381 samples inspected; the average small firm, 7.2. Each large firm is virtually certain of one or more samples per year; a small firm may hope to go uninspected for a year or more. (3) Most small firms buy the drugs in bulk form from large firms, and merely tablet and bottle the finished dosage forms. Pejorative comments by large firms, therefore, often imply at least some criticism of their own bulk drugs. (4) The United States Pharmacopoeia sets tolerance intervals above and below 100 per cent of labelled strength for the active ingredient in many drugs. If the lower limit is, for example, 95 per cent, it has been claimed that the producer might aim of 95 per cent strength and thereby achieve economies. For a small firm, however, the cost of goods sold may average about two-thirds of sales[42] and the cost of raw materials may constitute about half of the total cost of goods sold[43] so that raw materials costs may be in the neighborhood of one-third of total costs for a small firm. Hence a reduction of 5 per cent in raw material input would cut total cost by only about 1⅔ per cent, or by much less than it would require to begin to justify a procedure which would result in turning out a substantial number of tablets below the 95 per cent lower limit of tolerance.[44] (5) It is claimed that low priced sellers economize on quality control. Again, there seems to be little room for cost savings, for the cost of quality control is probably no more than 2½ to 3 per cent of the total of all costs[45]—surely not enough to justify a procedure which would make the violation of the drug laws a certainty.

Perhaps the best argument why quality differences might be expected to exist between higher and lower priced drugs is the contention that the Food and Drug Administration is understaffed and cannot make enough inspections. It is true that insufficient inspections are made with regard to both large and small firms. From January, 1950, to June, 1960, 7,699 samples of drugs produced by large firms ($10 million or more in annual sales) and 9,298 samples from small firms were inspected. During the same period, 84 incidents of irregularities developed in connection with drugs made by large firms, 79 of which were handled by the "drug recall" procedure and five of which led to legal actions. For small firms, 690 such irregularities developed, 206 of which involved drug recalls, and 484 of which led to legal actions. The ratio of legal actions is almost 100 to one, while the ratio of drug recalls is only five to two. Clearly, those irregularities involving large firms are much more frequently negotiated than those involving small firms.[46] For the large firms, there was a ratio of 1.1 percent of irregularities to total samples inspected. For the small firms, the ratio was 7.4 percent. The difference is a matter of degree; both figures are too high. The remedy is not to prohibit or to discourage by propagandistic activities the sales of drug at

---

[41] Data presented by Food and Drug Administration Commissioner G. P. Larrick, *id.*, pt. 22, at 12147.

[42] The only data available are 68 per cent for Panray Corporation. *Id.*, pt. 16, at 9375.

[43] The only data furnished by any one company for any product show a 52 per cent ratio for Carter's meprobamate. *Id.*, pt. 16, at 9157, 9161.

[44] The products of large firms commonly sell at prices as much as several hundred per cent above those charged by small firms, as has been indicated in the case of reserpine. Small firms can undercut large ones by economies in selling costs without any risk of violating the food and drug laws. Large firms spend about 25 per cent of sales revenue in sales promotion, according to data submitted to the Senate Subcommittee by the major firms during the hearings. It is clearly preferable to save 20 to 25 per cent of total cost by foregoing sales promotion in favor of price competition, than to save 1⅔ per cent in raw materials cost and risk a violation that might result in a fine large enough to put a very small firm out of business.

[45] Panray Corporation's quality control costs were 3 per cent of sales revenue; its profits before taxes were about 10 per cent of sales. Hence quality control for this small firm accounted for about 2.7 per cent of total cost. That this is adequate is evident from the fact that no samples of Panray products have been judged violative by Food and Drug Administration inspectors. Hearings on Administered Prices, pt. 16, at 9375, 9378.

[46] It must not be assumed that irregularities involving legal action were more serious, on the average, than those settled by drug recall. If anything, the reverse is true. The nature of 46 violations taken to court, for 32 large and small firms appearing in the record of the hearings, had to do almost entirely with deficiencies or excesses of the active ingredient; only three samples in 46 contained adventitious deleterious substances. Of 285 drug recalls, most were for lack or loss of potency, but several very serious irregularities appeared. For small firms, there were four instances of contaminated drugs, and one instance where the drug was suspected of causing agranulocytosis. For large firms, there were six instances of contaminated drugs, and one instance of suspected agranulocytosis, three instances where the drugs had caused liver damage or jaundice, one instance of induced acidosis, and one instance (again a recall, not a legal action) where polio vaccine had caused widespread paralytic poliomyelitis. *Id.*, pt. 22, at 12137–61.

lower prices. This would merely increase the monopoly power of the large firms in those few markets where price competition exists, and would therefore increase that part of the profits of the large firms which constitute monopoly returns. The obvious remedy is to provide adequate inspection. Food and Drug Administration Commissioner Larrick testified at the hearings that adequate ethical drugs inspection could be obtained at a cost of an additional $3,418,000.[47] For the 22 largest firms in the industry in 1958, net profits before taxes amounted to $562 million.[48] A simple computation may throw some light on the nature of the alternative. A reduction of net profits before taxes of $7.121 million would cut federal income tax receipts by 52 percent of this amount, or by $3.705 million. The net gain would therefore be the $7.121 million saved by drug customers, less the $3.703 million lost in tax receipts, or the needed $3.418 million. If adequate drug inspection could establish confidence in lower priced drugs to the extent that the resulting competition would lower major drug firm prices and profits by as little as 1.27 percent before taxes, the savings realized would pay for the expanded enforcement program. To this, the substantial benefits obtained by the elimination of inferior drugs must be added.

Nowhere is the policy of fostering imperfection of market information so much in evidence as in the controversy surrounding the issue of prescription by brand name instead of by generic name. Drug advertisements must, according to law, mention the appropriate generic name, but the generic name is universally given in much less prominent type face,[49] and is sometimes concealed by its appearance in an unlikely place in the advertising copy, or else is suppressed in favor of the full and formidable chemical nomenclature. These generic names do not lend themselves to practical use. Drug makers are free to designate the generic name for any new compound they market, and generally do so in such a way as to insure minimum use of the generic name in favor of the brand name.[50] To take a single example from among thousands, the compound with the brand name "Darvon" is burdened with the generic name dextropropophyphene hydrochloride.[51] Dr. Frederick H. Meyers of the University of California epitomized the predicament of the medical profession by commenting that ". . . unless these generic names are disciplined, the *United States Pharmacopoeia* will eventually become the dictionary of a nonsense language."[52]

Such generic names are at the very least inappropriate, since the generic name need serve no other purpose than simple identification of the compound. The definitive and adequately descriptive name is the chemical name, which is a verbal summary of the molecular structure, and which may considerably exceed in length even the most prolix examples of generic names. The generic name is usually merely an essentially unenlightening encumbrance, useless to chemist, pharmacist, and physician alike. An interesting contrast is provided by the procedure for designating new names for insecticides. Rules exist which require that names be short, distinctive, easily spelled, and in conformity with accepted scientific usage.[53] No such rules exist in the case of pharmaceutical preparations for human use.

The efforts made to suppress the use of generic names have been very effective. Medical schools teach doctors to prescribe by generic name.[54] In practice, physicians have proved vulnerable to the usual advertising appeals, and about 88 percent of all prescriptions are written in terms of brand names.[55] By this means, 88 percent of the prescription market is removed from the sphere of price competition.

---

[47] *Id.*, pt. 22, at 12132.

[48] Data submitted by firms to the Subcommittee. Profit estimate obtained by multiplying total sales by the 25.8 per cent weighted average sum of profits after taxes and taxes. Subcomm. Report 31.

[49] Dr. Austin Smith of the Pharmaceutical Manufacturers Association was at one point during the hearing unable to locate the generic name on an advertisement for a drug called "Dimetane" until given a magnifying glass by the Subcommittee counsel. Hearings on Administered Prices, pt. 19 at 10931.

[50] Drug-naming procedures are described in *id.*, pt. 21, at 11499, 11675, 11868.

[51] *Id.*, at pt. 21, at 11775–76.

[52] *Id.*, pt. 18, at 10401.

[53] Testimony of Dr. C. O. Wilson of Oregon State College, *id.*, pt. 21, at 11505.

[54] A survey sent to the 82 medical schools in the United States in 1960 brought 77 replies. Sixty-four schools teach only generic terminology; three teach generic and brand names together; and ten used brand names only under certain circumstances, such as the monopolization of a patented generic compound by a single firm. Subcomm. Report 226.

[55] 1958 National Prescription Survey, Pharmaceutical Extension Service, Rutgers University, as reported in Hearings on Administered Prices, pt. 15, at 8776.

The industry's reaction to the presence of Food and Drug Administration regulation has been apparent primarily in the area of the certification of new drug applications. In order to secure permission to market a new drug, the applicant must submit a sufficient amount of clinical and experimental data to establish that there is no significant danger connected with the proper use of the drug, *i.e.*, that its acute toxicity is sufficiently low to let it be considered a safe drug. It takes time to conduct sufficient experiments and carry out enough clinical studies to determine probable toxicity in general use. It takes still more time to have the application studied and processed. In drug marketing, many firms are often working on the same product at the same time, and each desires to cut the period between discovery and marketing to an absolute minimum, for the order of priority in market appearance usually determines the relative sales ranking for different brands of a given drug. Consequently, the motivation is to limit experimental and clinical work to the minimum acceptable level,[55] to skip stages in product development, such as the pilot-plant stage, and to influence the staff of the Food and Drug Administration in such a way as to facilitate rapid approval.[57]

The response of the ethical drug industry to its specific economic framework may be summarized as follows: the employment of the patent privilege to erect barriers to entry into the production of specific patentable drugs, the substitution of product differentiation for price competition, and the use of the accompanying techniques of sales promotion to minimize the impact of the price competition that might be offered by smaller firms. This is accomplished in three ways: the vast bulk of advertising done by the major firms tends, by its mere magnitude, to obscure the very existence of small, non-advertising generic sellers; the employment of opaque brand names for advertised drugs makes it formidably difficult for buyers to detect the existence of lower-priced, generic equivalents; and the campaign of disparagement renders suspect the quotation of a low price. No price competition need ever develop for patented drugs; for non-patented drugs, product differentiation, the adoption of deliberately confusing nomenclature, and the waging of a never-ending campaign of disparagement against low-priced drugs can effectively substitute for price competition, and prevent small, low-priced sellers from taking over any appreciable amount of the prescription drugs market, even though the absence of economies of scale or other barriers to entry will permit small sellers to undersell large by remarkable margins. This has been accomplished in the face of the presence of governmental inspection to insure the quality of all drugs merely by the extension of the policy of disparagement to include the adequacy of the Food and Drug Administration's facilities for making inspections.[58]

III. EVALUATION OF MARKET PERFORMANCE IN THE ETHICAL DRUGS INDUSTRY

A. Varieties of competition among drugs

There seem to be three main dimensions of competition: price competition, product competition, and product differentiation. While price competition is not

---

[56] This is naturally deplored by physicians and medical educators. Dr. C. D. Leake of Ohio State University noted pertinently: "There is no shortcut from chemical laboratory to clinic, except one that passes too close to the morgue." *Id.*, pt. 18, at 10418.

[57] Dr. Barbara Moulton, formerly with the Food and Drug Administration, testified that excessive deference was at times shown to impatient applicants. She testified that if the medical officer in charge of the evaluation of a new drug application is not satisfied as to the evidence of its safety, the applicant will frequently make an appointment with the medical director. She continued: "I have known such conferences to be followed by an order to the medical officer to make the new drug application effective, with the statement that the company in question has been evaluating new drugs much longer than the medical officer, and should, therefore, be in a much better position to judge their safety." *Id.*, addiction danger warning on the label of a tranquilizer was refused by her superior, whom pt. 22, at 12025. She related an experience of her own in which her request to place an addiction danger warning on the label of a tranquilizer was refused by her superior, whom she quotes as saying, "I will not have my policy of friendliness with industry interfered with." *Id.*, pt. 22 at 12032.

[58] The large firms, acting through trade associations, have been instrumental in contributing to the inadequacy of governmental inspection. Prior to 1953, the Food and Drug Administration had broad powers to inspect plants as well as products. The Factory Inspection Amendment of 1953, supported by the major firms, made it possible for drug makers to refuse to allow inspection of significant phases of drug operations. Large as well as small firms have not hesitated to avail themselves of this privilege. Testimony of G. P. Larrick, *id.*, pt. 22, at 12113.

apparent, closer study reveals that even in the case of sales to retail druggists, there is some price competition via quantity discounts and the shipment of "free goods" in excess of the invoice quantities ordered. In the hospital purchases markets, bids are frequently requested, and in some product lines, sharp price competition may develop.

Product competition proper is more easily achieved in the chemical industries than in some other fields. Research effort, whether domestic or foreign, public or private, may culminate in the discovery of a new compound with important therapeutic applications. If one firm is awarded the patent, other firms will attempt to modify the molecular structure of the compound in order to discover a "different" and patentable therapeutic agent which will in all probability be of use in the treatment of the same classes of disorders as the original drug, but which will hopefully be more potent, or less toxic, or will at least yield a different variety and incidence of concomitant side effects.

Product differentiation is seen in its purest form when several firms obtain licenses from the patent holder and then proceed to market the identical compound under different names, each of which must then be intensively advertised. The choice of a name for such a drug constitutes an exercise in product differentiation at a higher level than is usually encountered in the commodity market, since the name typically abandons or repudiates description of the good and refers instead to nothing outside itself, stressing only its abstract (and presumably unique) identity.[59] Product differentiation efforts are of course also enlisted in behalf of products which are physically different from each other. Since "ethical" drugs are ethical in the sense that they cannot be bought over the counter or advertised to the public at large, the market consists of the private physicians and hospital pharmacists who can order them. This is a relatively small and well-defined market which can be intensely saturated by advertisements and the employment of itinerant salesmen. There is, in fact, some evidence that the intensity itself, at least in some quarter.[60]

The role of research as the driving force behind both the search for new products and the devising of minor molecular modifications on old ones should be considered separately at some length. Lack of space, however, allows only a brief summary. Spokesmen for the major drug firms invariably defend the height of their prices, or the gap between computed costs and market prices, by reference to the vast sums spent on research in the interests of advancing the cause of health and medical science. It is also asserted that in the absence of the patent incentive, research would disappear. Both statements may be questioned. It may be doubted that they can be reconciled. It is probably difficult to convince an economist that the primary purpose of drug research is not to increase profits. This is economically desirable only if no cheaper way can be found to insure an adequate supply of new and improved drugs. That research is the monopoly of the large and profitable firms, or that the most profitable firms do most of the research, may be questioned. That research outlays are a major or even a very significant factor in the total cost picture for most major drug firms may further be questioned. Finally, that the patent privilege is a necessary incentive to elicit truly productive drug research may be disputed. As a brief summary of the evidence in regard to these issues, the following facts must suffice:

1. The size of the research budget as a per cent of the sales dollar varies inversely with the size of the firm. For a group of 22 large firms and two representative smaller firms, the most profitable firm (Carter, with a profit before taxes of 43.8 per cent of sales) had the smallest research budget (2.7 percent of sales), while the least profitable firm (Panray, 10 per cent of sales) is estimated to have had the highest relative research budget (15 per cent of sales), during the year 1959.

---

[59] An example may make the contrast with the usual type of product differentiation by brand name more apparent. Dr. Solomon Garb of the Albany Medical College of Union University submitted that if the drug makers took over the manufacture of canned beans, then rather than selling "Pfizer's Beans" or "Parke, Davis and Co. Beans" they would prefer to coin novel and unique anagrams such as "Sneabs" or "Nabes," or adopt undescriptive slogans like "Lo Cals" or "Hi Pro's" and abandon entirely the generic noun. *Id.*, pt. 18, at 10481.

[60] See particularly the testimony of Drs. Bowes, Console, Seidell, Weinstein, Bean, Garb, Leake, and Meyers, *id.*, pt. 18.

2. For the 22 largest drug firms in 1959, research outlays comprise the smallest part of the budget. Of the sales dollar, 32.1 cents represented cost of goods sold, 25.8 cents was net profits before taxes, 24.8 cents went for selling costs, 11.0 cents for general and administrative expenses, and only 6.3 cents for research.[61] Selling costs average about 394 per cent of research costs. If the cost of drugs is excessive, it is not because of the magnitude of the research budget.

3. The growth of a large American drug firm has typically depended but little upon the quality of its own research. The largest firm Pfizer did not begin to grow rapidly until 1941, when it was 92 years old, and then its growth was largely in the nature of a windfall gain. Pfizer was one of only a few firms with experience in fermentation methods when wartime demand arose for fermentation-produced penicillin (discovered in England). The same is true for Bristol, which owes its present position largely to the purchase in 1943 of a small firm with some experience in fermentation techniques.[62] Smith, Kline, and French owes its growth not to its own research, but to its ability to take advantage of promising Europeon discoveries. To a lesser extent, this is true of Carter. Additional examples could readily be given.

4. More basically new drugs have been found in European countries without drug product patents than in the United States with its unrestricted drug patent privileges. On the basis of an inclusive list (including molecular modifications as well as basic new discoveries) drawn up by the Senate Subcommittee staff and revised by the Pharmaceutical Manufacturers' Association, since 1886, 82 drug discoveries have been made in countries without product patents, compared with 79 in the United States, only 60 of which were found in the laboratories of drug firms. Fifteen drugs were found in foreign countries with drug patents.[63] Hence, 75 drugs were found by American commercial firms and in foreign countries with product patents, while 101 drugs have been found in countries without product patents, and by noncommercial American investigators.

5. Most American drug firm research is not basic research and is not very fruitful.[64] With regard to four drug categories, American industry is responsible for the basic discovery in only one area, cortisone in the corticosteroid field. In oral antidiabetic drugs, the basic drug was found in Germany; in tranquilizers in France, Switzerland, and England; and in antibiotics, in England. Minor modifications have been made by American firms in all four areas, for it is in this sort of research (or more properly, product development) that the domestic industry excels: the further exploitation of an already accomplished basic discovery. It is in this sort of product development-research that the risks are least.[65] Such activities in this country have been fruitful, as in the field of broad-spectrum antibiotics, where important new drugs were found, but no basic research was involved. The basic discoveries had been made in England in 1929, the problems of large-scale production had been solved by government scientists during the second world war, and the observation was current that many molds occurred in nature, some of which might be capable of yielding new antibiotics. All that remained to be done was to collect and screen soil samples, a tedious, trial-and-error procedure of a routine nature, which the technicians involved have

[61] Subcomm. Report 31.
[62] Hearings on Administered Prices, pt. 24, at 13825.
[63] Subcomm. Report 116–18.
[64] Dr. H. J. Weinstein, former medical director of Pfizer, testified that much drug firm research "has little relationship to what most people engaged in academic and research activities would consider to be scientific research." He gave as examples the practices of molecular manipulation, the aimless proliferation of combinations of existing drugs which lack flexibility and compound the problems of dosage and toxicity, and the "battle of the additives" where much energy is devoted to showing that the addition of certain substances to an existing drug will intensify its activity. Hearings on Administered Prices, pt. 18, 10243–44. Dr. Console testified: "I think the majority of it [research] is in that category . . . with many of these products, it is clear while they are on the drawing board that they promise no utility. They promise sales." *Id.*, pt. 18, at 10380.
[65] Dr. A. D. Console, former medical director of Squibb, testified that "they [the drug firms] stress that there are many failures for each successful drug. This is true since it is the very essence of research. The problem arises out of the fact that they market so many of their failures." He continued, "I doubt that there are many other industries in which research is so free of risks . . . with a little luck, proper timing, and a good promotion program, a bag of asafetida with a unique chemical side chain can be made to look like a wonder drug." *Id.*, pt. 18, at 10372–373.

sometimes referred to as "making glorified mud-pies." [66] Furthermore, since no basic research was involved, the returns to more and more intensive exploitation of the same method have approached zero. No important antibiotic has been found in the United States since 1955, and the same is true in other drug fields.

B. Degree of product substitutability among drugs

With this background, which cannot pretend to do justice to the many complexities of the research issue, it is possible to turn to the evaluation of market performance in the four drug product categories investigated during the Senate Subcommittee Hearings, corticosteroid hormones, tranquilizers, oral antidiabetic drugs, and antibiotics. These four categories accounted for about 42 per cent of all ethical drug sales in 1959. The corticosteroids market will first be described, and will then be used as the point of departure for parallel comparisons with other drug product categories. By virtue of the emphasis which it places upon the continuous obsolescence of existing corticosteroids through the development of new synthetic steroid hormones eligible for patent protection, this market may be taken as typical of ethical drugs as a whole.

The important corticosteroids in 1959 were seven in number: cortisone, hydrocortisone, prednisone, prednisolene, methylprednisolone, triamicinolone, and dexamethasone. All of the corticosteroids are to some extent substitutes for one another in the treatment of many disorders. Striking evidence of this is available on the basis of a 1959 study by the Pharmaceutical Manufacturers Association, showing that all seven of the major corticosteroids were prescribed in every one of the seven disease groups requiring such therapy, with but two exceptions.[67] No one of these drugs is apparently characterized by advantages so great that its actual use is always exclusively in order in a certain disease group. Obviously, for individual diseases or for individual classes of patients this may not be the case, for the disease groups are not very narrowly defined, and even within the limits of a single disease, patients of different ages or with different degrees of proneness to complications might require different specific drugs. Nevertheless, it is striking that in no instance is as much as two-fifths of any disease group treated by a single drug. In such a situation it may well be the case that no drug has any commanding advantage over all the others for any major disease group; hence the patterns of demand may be greatly influenced by sales efforts, and the products may be differentiated in terms of their claimed freedom from side effects, their relative potency, their probable impact on certain types of patients, and in fact upon anything but their relative prices.

Similar evidence is available for the tranquilizers market from the same source. This time, each of the seven major tranquilizers is prescribed in every one of the five disease groups requiring such therapy.[68] The maximum usage in any disease group is 41.4 per cent. If all drugs were prescribed at random for all disease groups, each would be used in 14.3 per cent of all treatments. In only four instances is the frequency of use more than twice what one would expect on the basis of random choice. Substitution among drugs is apparently as much the rule in tranquilizers as in corticosteroids.

The oral antidiabetics drug market is more readily described. There are three such drugs, tolbutamide, chloropropamide and phenformin. All seem to be fairly close substitutes, and tolbutamide (Upjohn's "Orinase") and chlorpropamide (Pfizer's "Diabinese") have been subjected to strenuous competitive product differentiation programs. The market shares follow the order of priority of market appearance, the earliest to appear having the largest share. These products may not be perfect substitutes, however, for it has been claimed that chlorpropamide is effective in some cases where tolbutamide is not,[69] and there is evidence to show that side effects are more frequent and severe with chlorpropamide.[70]

It is impossible to summarize the extent of possible substitution among antibiotics in brief compass. The main areas in which substitution is possible are among the five "broad spectrum" antibiotics, all of which are considered interchangeable in many uses, and all of which are highly advertised; and between

[66] Subcomm. Report 124.
[67] Hearings on Administered Prices, pt. 19, at 10826. (One drug was prescribed in only five groups.)
[68] *Id.*, pt. 19, at 10827.
[69] Testimony of Dr. C. J. Donovan, formerly of Upjohn, *id.*, pt. 20, at 11070–72.
[70] Testimony of Dr. Henry Dolger, Mount Sinai Hospital of New York, *id.*, pt. 20, at 11146.

penicillin, which is unpatented and not highly advertised, and the broad spectrum antibiotics. Within its range of activity, penicillin is usually more potent than the broad spectrum drugs.[71] But micro-organisms tend to become resistant to certain antibiotics; many of the newer antibiotics are said to be effective against strains resistant to the older ones. Among the broad spectrum drugs, chloramphenical may currently be regarded as somewhat distinct from the others, in that fewer strains have become resistant to it, and for two reasons. First, the use of chloramphenicol was discouraged to some extent in the early 1950's when publicity was given to the incidence of its lethal side-effects; hence its less frequent use during that period provided less occasion for the development of resistant strains. Second, this drug resembles penicillin more than do the other broad-spectrum antibiotics in that microorganisms do not easily develop resistance to it;[72] chloramphenicol is, in fact, the most lethal of the broad spectrum antibiotics to micro-organism and patient alike.[73] The incidence of side effects is often more important than the spectrum of primary efficacy of the antibiotic as far as drug demand is concerned. Advertising may influence prescription demand by stressing the alleged absence of side effects as well as therapeutic efficacy. The absence of any advertising may conversely prevent the wide-scale use of extremely efficacious antibiotics. Bacitracin is a case in point. The drug is so free of side effects that some of its preparations may be bought over the counter. The patent is owned by the federal government, and royalty-free licenses may be obtained. The only ethical drug firm currently producing this drug, Pfizer, also sells its own patented antibiotics, tetracycline and oxytetracycline, and does no sales promotion for bacitracin, save of a negative sort. Originally the Food and Drug Administration restricted its use to hospitals, but later removed the restriction, but as of September, 1960, Pfizer (despite protests from the discoverer of the drug) continued to print the former restriction on bacitracin labels and brochures.[74]

C. Product competition in individual drug markets

The earliest of the corticosteroids, cortisone, was first synthesized in 1944 by Merck, and commercial production began in 1948.[75] The drug first proved its effectiveness in relieving the symptoms of rheumatoid arthritis and related clinical syndromes, although concomitant side effects presented complications. Cortisone is a naturally occurring substance in the human body (adrenal cortex) and no product patent could be obtained. Process patents presented no real barrier to entry, and Merck soon had competition from a number of large and small firms.[76]

Squibb discovered fludrocortisone in 1953, by modifying the hydrocortisone molecule. Squibb obtained a patent, sells the drug under the name "Florinef" and has apparently licensed no other sellers. More potent than its parent, its side effects are also more severe.[77]

Prednisone was the next corticosteroid to be found, along with the related compound prednisolone. These discoveries are credited to the Syntex Corporation of Mexico in 1955. By nature still further modifications of the cortisone molecule, they are generally felt to constitute some actual improvement over the parent substance. They are used to treat the same disorders, and have about five times the antirheumatic activity of the parent steroids.[78] No patent has yet been awarded for either drug, and several large firms and many small firms sell both drugs.

Some medical authorities hold that no essential advances have been made in the corticosteroids field since prednisone and prenisolone were introduced. Upjohn found methylprednisolone in 1957 as a further modification of the pred-

---

[71] L. S. Goodman & E. Gillman, The Pharmacological Basis of Therapeutics 1322–23 (1956).

[72] Testimony of Dr. Maxwell Finland of Harvard University, Hearings on Administered Prices, pt. 24, 13947–49.

[73] In a nation-wide survey of all severe antibiotic reactions from late 1953 to early 1957, 25 deaths from severe skin reactions and blood duscrasias were associated with the use of chloramphenicol, and only two deaths from the same causes were associated with the use of the tetracyclines. L. Meyler, Side Effects of Drugs 129–30 (1960).

[74] Testimony of Dr. F. L. Meleney of the University of Miami, Hearings on Administered Prices, pt. 24, at 14195.

[75] Testimony of Dr. E. C. Kendall of Merck, *id.*, pt. 14, at 8018–20.

[76] Testimony of J. T. Connor, president of Merck, *id.*, pt. 14, at 8025–26.

[77] Modern Drug Encyclopedia and Therapeutic Index 471 (E. P. Jordan ed. 1958).

[78] Meyler, *op. cit. supra* note 73, at 148.

nisolone molecule, and claimed that it had fewer side effects than other corticosteroids, while having the same general field of use. Upjohn obtained the patent and then licensed Schering to produce the drug.

Lederle brought out triamcinolone in 1958, another molecular modification of cortisone, with a potency about equal to that of methylprednisolone, with the same general therapeutic applications as the other steroids, and with the same claim of fewer side effects. Lederle obtained the patent, and licensed only Squibb.

The latest modification of the cortisone molecule to be marketed is dexamethasone, brought out by Merck and Schering in 1958. This drug is said to be the most potent of the corticosteriods, being about seven times as potent as prednisone,[79] and is also claimed to have fewer side effects. No patent has yet been awarded. Merck, Schering, and CIBA were producing this drug in 1959.

A Survey of worldwide experience with these newer drugs prompted the opinion that their side effects are ". . . certainly no less than those of prednisone and prednisolone. . . . With the newer drugs, digestive accidents, fractures, adrenocortical inhibition, and indfections are still the main dangers. In contrast to the manufacturers' statements, risks are in no way eliminated by the new derivatives." [80] Since it takes years to determine the true incidence of side effects for a given drug, medical opinion is as yet inconclusive, but the most frequently expressed view is that all the corticosteroids will prove to have the same side effects in the long run.[81] Thus product improvement has here taken the form of product competition by means of trial-and-error molecular manipulation, and product differentiation by means of massive selling efforts, rather than by revolutionary advances which would make all previous corticosteroids obsolete.

In other drug markets, the development of new products also relies heavily upon the devising of molecular variations on former products. In oral anti-diabetic drugs, tolbutamide was found in Germany and licensed by Upjohn for sale in the United States as "Orinase." Upjohn carried out extensive clinical testing, and the drug was released for general prescription use in the United States in June, 1957. Other firms conducted research in order to produce a patentable substitute compound. Lilly and Pfizer came up with chlorpropamide, a slight molecular modification of tolbutamide, and Pfizer began to market this product as "Diabinese" in November, 1958. Lilly withdrew its patent application in favor of Pfizer; at present no patent has yet been awarded. United States Vitamin and Pharmaceutical Corporation later discovered phenformin, a molecular modification of certain compounds subjected to study by many scientists prior to the discovery of insulin in 1920, and marketed it as "DBI." [82] Tolbutamide appears to be unusually free from side effects. Chlorpropamide, on the contrary, is not; nor is phenformin.[83] Some physicians are of the opinion that advertising, rather than therapeutic superiority, is responsible for the sales of the later oral antidiabetic drugs.

In the tranquilizers market, the basic research on the more potent tranquilizers was done largely in France, with the development of certain phenothiazine compounds with sedative effects. Smith, Kline and French was licensed to sell chlorpromazine ("Thorazine") and prochlorperazine ("Compazine") in the United States. "Thorazine" was released for sale in 1954 as an anti-emetic; the tranquilizer properties were advertised only later when it was found that a market existed in the mental institution area. "Compazine" was introduced in 1956 as a tranquilizer, and is a phenothiazine compound similar to chlorpromazine. The market for tranquilizers grew, and so did the number of such drugs. Promazine was marketed as "Sparine" by Wyeth in 1957; again the drug is a molecular modification. Amusingly, Wyeth began to advertise promethazine hydrochloride ("Phenergan"), an antihistamine originating in France, as a tranquilizer.[84]

---

[79] Report from Pharmaceutical Manufacturers Association, Hearings on Administered Prices, pt. 19, at 10845.

[80] Meyler, *op. cit. supra* note 73, at 148.

[81] Testimony of Dr. Russell Cecil, consulting medical director of the Arthritis and Rheumatism Foundation, Hearings on Administered Prices, pt. 14, at 7985.

[82] Subcomm. Report 123.

[83] The use of chlorpropamide is associated with liver damage and other mischief. Dr. Henry Dolger, in testifying at the Hearings, concluded: "I deplore the premature prosecution of unnecessary and inadequately investigated modification of sound drugs and their indiscriminate and unsubstantiated promotional campaigns." Hearings on Administered Prices, pt. 20, at 11149.

[84] In 1952, 20 per cent drowsiness was noted as a side effect. In 1959, it was stated in an advertisement that "Phenergan relieves apprehension, relaxes the patient, and produces light sleep." *Id.*, pt. 16, at 9272.

Other molecular modifications followed: perphenazine ("Trilafon"), trifluoperazine ("Stelazine"), and trimepazine ("Temaril").[85] In the mild tranquilizers field, a Dr. Berger had discovered mephenesin, a muscular relaxant, in England, and came to the United States to patent a molecular modification, meprobamate, as a tranquilizer, which was marketed by Carter ("Miltown") and Wyeth ("Equanil") in 1955. Both products were highly advertised, a condition of Wyeth's license from Carter being that Wyeth would spend at least 20 percent of total meprobamate sales revenue in promoting its sale.[86] Smith, Kline and French found itself with the potent tranquilizers and no mild tranquilizer; it then began to advertise "Compazine" as a mild tranquilizer. "Compazine" is actually much more potent than "Thorazine."[87] Medical opinion inclines to the view that none of the later modifications of the original phenothiazines has been a marked improvement. Dr. Lehmann testified: "There hasn't been a very much better one than the very first ones that came out, in the six or seven years of frantic research since then."[88] No diminution in the incidence of side effects could be demonstrated.[89]

The situation in the antibiotics market is largely the same, but of greater complexity. The discovery of penicillin in England in 1929, and the further discovery of streptomycin at Rutgers University in 1943, demonstrated the possibility that many naturally occurring molds might have antibiotic properties. Many drug makers set themselves the task of the trial and error screening of thousands of soil samples. The first success was encountered at Yale University in 1947 under a Parke, Davis grant, the antibiotic chloramphenicol ("Chloromycetin") being isolated. Lederle next discovered chlortetracycline in 1948, and named it "Aureomycin." Pfizer then found oxytetracycline in 1949, and designated it "Terramycin." Pfizer managed to elucidate the chemical structure of oxytetracycline, and with the molecular structure known, molecular manipulation became possible in antibiotics, although to a more limited extent than is possible with drugs which can be made synthetically. In 1955, Pfizer developed tetracycline as a molecular modification of chlortetracycline.[90] The screening of soil samples continued. Lilly in 1952 discovered erythromycin. Pfizer in 1953 brought out a closely related analog, carbomycin, which affected the same bacteria as erythromycin, but which proved less effective in human disease than in the test tube, and which was finally withdrawn from the market in 1960.[91] In 1956, Pfizer introduced another closely related analog of erythromycin, oleandomycin; and in 1957, a modification of oleandomycin, triacetyloleandomycin, which was advertised as a great advance, since the same oral dose as in oleandomycin produced somewhat higher concentrations of the drug in the blood. To offset this, Lilly in 1958 modified its original erythromycin to market it in the form of its propronyl salt, which is said to produce a higher concentration of this drug in the blood than triacetyloleandomycin.[92] Hence, five drugs were produced to serve the purpose of one (since slightly higher doses of erythromycin could serve the same purpose as the newer derivatives). Another broad spectrum antibiotic was brought out by Lederle in 1959, demethylchlortetracycline, or "Declomycin," a slight modification of Lederle's chlortetracycline. Fewer side effects than other broad spectrum antibiotics were claimed. Not all physicians agree.[93] The later penicillins are patentable variations of the earlier original penicillin which the British discoverer did not bother to patent. In 1952, four American firms made a substitution of benazthine for procaine in the penicillin compound; the parties

---

[85] Subcomm. Report 121.
[86] Testimony of Irene Till, Hearings on Administered Prices, pt. 16, at 9208.
[87] Testimony of Dr. Heinz Lehmann of Verdun Protestant Hospital, Montreal, *id.*, pt. 16, at 9068–73.
[88] *Id.*, pt. 16, at 9029.
[89] Physicians employed by drug firms who testified sometimes found it easier to ignore side effects than one suspects their drug users could. Dr. Gibson of Merck went so far as to introduce his aesthetic preference for facial contours as the determinative criterion for evaluating the side effect of "moon face" encountered in steroid therapy. In commenting upon this side effect (caused by redundant retention of cellular fluids, a circumstance conducive to additional strain on the cardiovascular system), Dr. Gibson testified: "When those people [malnourished patients] got a moonface I think it is a desirable effect." *Id.*, pt. 14, at 8188.
[90] Subcomm. Report 116–17, 124.
[91] Testimony of Dr. H. F. Dowling of the University of Illinois, Hearings on Administered Prices, pt. 24, at 14167.
[92] *Id.*, pt. 24, at 14167–68.
[93] Several of the physicians who testified during the Hearings professed to see little difference among the broad spectrum of antibiotics. See *id.*, pt. 18.

agreed to concede priority to Wyeth in return for suitable cross-licensing assurances. In 1951, Lilly secured a patent which embraced phenoxymethyl penicillin among other compounds. Lilly apparently did not recognize its value; not until an Austrian firm applied for an American patent on it did Lilly learn of its potential. Lilly then entered into cross-licensing agreements with that firm, and did not market its product until 1955, four years after discovery. The next penicillin, phenoxyethyl penicillin, is a slight modification of phenoxymethyl penicillin, was found in England, and is marketed under four different brand names in the United States.[94]

D. Price competition in individual drug markets

In general, the presence or absence of price competition in drugs depends upon the presence or absence of patent protection. Even where some price competition exists, however, it is rendered all but ineffective in the prescription market because of the inability of the lower price of the smaller producer to overcome the disadvantages of the obscurity inherent in small size and the disparagement associated with low price. In such cases, effective price competition is limited to the hospital formulary and government bid markets. Unpatented drugs compete with patented drugs in two out of the four markets under consideration, there being no important unpatented tranquilizers or oral antidiabetic drugs.

In the corticosteroid hormones market, cortisone and hydrocortisone were ruled ineligible for patent protection, and no patents have yet been issued for prednisone, prednisolone, and dexamethasone; all other corticosteroids have been patented. For cortisone, the price was $20 per gram in October, 1951, but intense price competition forced the price down to $5.48 per gram in 1954, at which level it has since remained constant.[95] But the case of prednisone is particularly instructive and merits review at some length. Five major drug firms filed patent applications for prednisone, and an interference proceeding was declared by the Patent Office. The parties involved undertook to make interim arrangements among themselves, involving the payment of "interim royalties" to one of the firms, and as a result all five firms began selling prednisone at the identical price of $17.90 per bottle of one hundred 5 milligram tablets. The Syntex Company of Mexico later became a party to the interference proceedings but was not allowed to participate in the interim arrangements because of its known reputation for price competition.[96] Syntex then retaliated by selling the drug in bulk form to small drug firms in the United States, and two of the major firms met its lower price for bulk sales. Syntex then cut prices still lower, from $10.01 per gram in 1957 to $2.36 per gram in late 1959.[97] The major firms' prices at wholesale for the finished drug remained constant at $17.90 per bottle, but small firms were selling at prices as low as $1.75 per bottle by 1959.[98] The lower-price brands, however, were unable to obtain more than eight or nine per cent of the prescription market. In the absence of extremely imperfect market information, such great differences in price would be impossible. The larger firms apparently found it much more congenial to disparage the products of the lower priced sellers than to meet their competition.

Had they cut prices, however, there would have been a great deal of room for price competition. The actual full cost of production for a bottle of one thousand 5 milligram tablets of prednisone is only $8.99.[99] This cost is but 5.07 per cent of the equivalen wholesale price of $179.00.

No price competition whatsoever developed among the patented corticosteroid hormones. Drugs with increasing potencies per gram appeared from year to year, but the price per tablet of each new drug was set so as to achieve a price per equivalent therapeutic dosage unit identical with those of earlier drugs with different dosage units per tablet. In 1958, however, two firms applied for patents upon dexamethasone and marketed it (while patent interference proceedings were still pending) at a price ten per cent below that being charged for its pat-

[94] Subcomm. Report 124–25.
[95] Hearings on Administered Prices, pt. 14, at 7884.
[96] *Id.*, pt. 14, at 7920.
[97] *Id.*, pt. 14, at 8045.
[98] Subcomm. Report 17.
[99] Hearings on the Drug Industry Antitrust Act before the Antitrust Subcommittee of the House Comm. on the Judiciary, 87th Cong., 2d Sess. 15 (1962). (Hereinafter cited as "Hearings on Drug Industry Antitrust Act.")

ented substitutes.[100] No small firms entered production, however, because of the higher costs involved than in the case of prednisone. Here, lower prices could be quoted by major firms in spite of appreciably higher costs because of the very large profit margins made on substitute patented drugs.

In the oral antidiabetic drugs market, no comparisons between the prices of patented and unpatented drugs can be made in view of the absence of the latter. Costs may, however, be compared with prices in the case of "Orinase" (Upjohn's name for tolbutamide). The Senate Subcommittee staff estimated the cost of producing one thousand 500 milligram tablets to be $6.86, which is only 8.0 per cent of the wholesale price of $83.40. Upjohn, however pays a royalty of 7.5 per cent of the sales price to the German patent holder, which increases the cost to 15.5 per cent of the wholesale price.[101]

None of the tranquilizers is unpatented, but price competition has been allowed to develop in the case of a single drug, reserpine. CIBA obtained the patent, but elected to take the unique course of licensing the patent widely, and also allowing the licensees to sell the drug in bulk powder form.[102] The licensees sold the bulk powder to small firms which fabricated their own capsules and proceeded to cut prices by as much as 90 per cent below CIBA's price. CIBA charged a wholesale price of $39.90 per bottle of one thousand 25 milligram tablets; other firms quoted prices as low a Panray's $2.65.[103] The cost of production of such a bottle is a matter of record: sixty-three cents.[104] This is only 1.6 per cent of the wholesale price.

Price competition in the antibiotics market is brisk for unpatented drugs, such as the earlier penicillins, and for streptomycin, the patent for which was licensed by Merck to several other firms. Price competition is, however, extremely sluggish for the later patented antibiotics.

Penicillin declined in price by about 85 per cent between 1945 and 1948, and by almost 90 per cent between 1949 and 1952.[105] Between 1952 and 1960 the price fell by another 80 per cent, although there was a marked increase from 1955 to 1957. Increasingly efficient production methods, coupled with the persisting competitive pressures resulting from free entry, made possible such a decrease in the price of this extremely important drug. Streptomycin prices declined by 95 per cent, 60 per cent, and 88 per cent for the same intervals, or by about the same amount as penicillin, and for the same reasons.

Broad spectrum prices show very different patterns. Lederle set an initial price of $15.00 for sixteen 250-milligram capsules of chlortetracycline. This may have been above the profit maximizing price; within two months it was cut to $10.00, the same price at which Parke, Davis entered the market in March. 1949. On February 1, 1950, both sellers cut prices by 20 per cent, to $8.00. In April, 1950, Pfizer entered the market with oxytetracycline, selling at $8.40, 5 per cent above its rivals. In May, 1950, these rivals cut their prices by 25 per cent, to $6.00, but Pfizer did not meet the cut until November. On September 27, 1951, Pfizer cut its price by 15 percent to $5.10, and by November 1 its rivals had met the price cut. Tetracycline and demethylchlortetracycline came later, and sold also for $5.10. No further price cuts were made for almost nine years, when Pfizer in August, 1960, again acted as the price leader by initiating a 15 percent "trade discount" adjustment which eventually became a 15 per cent price cut to $4.34.[106] (This price cut came one month before the Senate Hearings on antibiotics were scheduled.)[107] From 1948 to 1949, broad spectrum prices fell by one-third; from 1949 to 1951, by 49 per cent, and from 1951 to August, 1960, by zero per cent. It

---

[100] Hearings on Administered Prices, pt. 14, at 8020.

[101] Subcomm. Report 20.

[102] Reserpine is the name given to the active substance isolated from the root of Rauwolfia Sepentina, a species of tropical plant the medicinal properties of which had been known in India for centuries. CIBA was therefore perhaps uncertain as to the validity of its patent.

[103] Hearings on Administered Prices, pt. 16, at 9369.

[104] Hearings on Drug Industry Antitrust Act 16. (Cost data from McKesson and Robbins.)

[105] FTC, Economic Report on Antibiotics Manufacture 164, 179–80 (1958).

[106] Hearings on Administered Prices, pt. 24, at 13664.

[107] Mr. Duncan of Lederle may have failed to sense the humor of the situation at one point during the Hearings. He testified that Pfizer had cut its prices on a Saturday in order to "steal a march" on the industry. The following colloquy ensued:

"Senator KEFAUVER: You mean after ten years of operation, they would suddenly steal a march on you?"

"Mr. DUNCAN: Yes, Senator. These things happen very rapidly. If you can get any kind of an advantage on your competitors, you try to do so." (Hearings on Administered Prices, pt. 24, at 13728.)

can be maintained that since 1951 the costs of producing broad-spectrum antibiotics have fallen by about as much as the cost of producing penicillin, since the production methods employed are broadly the same.[108] A decline of 90 per cent in penicillin prices since November, 1951 is to be contrasted with a 15 per cent decline for broad spectrum antibiotics. Even if broad spectrum costs fell only half as much as penicillin costs, the burden on the drug consumer due to the absence of pure competition in broad spectrum antibiotics would be an extra 30 percentage points in price.

The Senate Subcommittee staff made production cost estimates for chloramphenicol, and made public Bristol's tetracycline production costs. It cost Bristol $1.67 to produce a bottle of one hundred 250 milligram tablets which sold at wholesale for $30.60. Production costs were therefore 5.5 percent of the wholesale price.[109] Bristol paid a royalty of $1.03 to Pfizer, increasing Bristol's costs to 8.9 percent of the wholesale price.[110] The estimated cost of producing a bottle of one hundred 250 milligram capsules of chloramphenicol (the production and sale of which is monopolized by Parke, Davis) is $1.52, or only 5.0 percent of the wholesale price of $30.60.[111]

Limitations on space prevent adequate treatment of antibiotics price competition in the government bid market. For unpatented drugs, competitive bidding has reduced prices steadily. For the broad spectrum antibiotics, only sporadic competition was manifest until foreign bids were accepted for tetracycline at less than half domestic bid levels. Price competition was absent until foreign firms, not encumbered by restrictive patent agreements, were allowed to bid.[112]

IV. CONCLUSIONS

It seems evident, particularly on the basis of experience in the corticosteroids, tranquilizers, and antibiotics markets, that genuine price competition among ethical drugs is effectively prevented, for the most part, by the existence of product patent privileges. Patent or patent application holders may exercise restrictions on output, and the resulting high level of prices used to finance selling campaigns which contribute to the otherwise serious imperfections of market information and make it impossible for small sellers of generic name products to obtain any significant share of the retail prescription market, even for those products not (or not yet) protected by patents.

It is possible that the ethical drug industry might be highly competitive, perhaps almost purely competitive, in the absence of patent protection and the brand name monopoly or oligopoly it makes possible. The only important economy of scale in most drug markets seems to be that of large scale advertising and distribution. The extent to which the costs of financing a sizable selling effort can increase the price of a product is well illustrated by the experience of the Panray Corporation in attempting to market its own brand of the tranquilizer reserpine, which is highly advertised by CIBA under the brand name "Serpasil" and is sold by CIBA at a wholesale price of $39.50 per bottle of one thousand 25-milligram tables. Panray found it impossible to sell to druggists for less than $20–$21 under its own brand name, "Serpanray." After it abandoned its selling efforts, it was able to make a profit by selling reserpine under its generic name at $2.65 per 1,000 tablets.[113]

Simple yet radical remedies are required for the institution of competition in the industry. The first step is the abolition of the patent privilege as it applies to drug products. Most of the countries in the world with patent laws grant no patents on drug products as a matter of public policy. Provisions should be made for the compulsory licensing, at reasonable royalty rates, of drug process patents. It is often claimed that such measures would destroy the incentive necessary for

---

[108] Subcomm. Report 82. Chloramphenicol costs have probably declined by still more since Parke, Davis discovered a way of producing this drug synthetically.

[109] Hearings on Drug Industry Antitrust Act 17.

[110] Hearings on Administered Prices, pt. 24, at 13876.

[111] *Id.*, pt. 24, at 13971–78.

[112] The hypothesis of the absence of real price competition is afforded some support by the testimony of Admiral Knickerbocker. The decision to seek foreign bids was made following a meeting of Military Medical Supply Agency officials with Pfizer representatives, during which meeting, according to Admiral Knickerbocker's testimony, "Mr. Cooney [Pfizer's sales representative] in rather an unguarded moment, made the statement that the price of tetracyline would stay where it was until Pfizer did something about it." *Id.*, pt. 24, at 13819. Had all firms been yoked by patent restricions, his posiion would have been stronger.

[113] Testimony of Myron Pantzer of Panray, *id.*, pt. 16, at 9366–68.

pharmaceutical research. A major shift of pharmaceutical research endeavor from private firms to public, university, and foundation channels could scarcely avoid resulting in gains. Private firms carry on relatively little basic research, and most of their so-called research budgets would probably better be described as applied research, product development, and product promotion. A large proportion of some very scarce human resources is wasted in the competitive duplication of research programs carried on concurrently by virtually all the major firms as is attested to by the virtually simultaneous discovery of many of the corticosteroids and antibiotics. In the absence of economies of scale, the removal of patent protection would induce new entry of many firms producing under constant cost conditions. The best remedy for a vertical or nearly vertical industry demand curve is a horizontal industry supply curve.

The second step is to make Food and Drug Administration inspection and regulation more adequate to meet the needs of the drug buyer and prevent his exploitation. The Food and Drug Administration should be given sufficient additional funds to allow it to inspect adequately the operations and products of all drug makers, large and small. It should further be given authority and additional funds requisite for the control of all drug advertising, and the labelling laws with respect to drugs should be amended to require that all ethical drugs be identified solely by the generic name, in conjunction with the name of the seller employed as an adjectival modifier. All advertising as well as labelling should be undertaken with respect to the generic name, coupled with the name of the seller, rather than concentrating upon the promotion of quasi-chemical syllable assortments. All advertisements should prominently feature suggested retail prices.

Both measures would tend to promote the same ends. The abolition of product patents would lead to the elimination of the monopoly profits which are necessary to finance intensive selling efforts and wasteful product-modification oriented research, and the departure of such profit possibilities would lessen the need for such advertising and the motivation for such research. The reform of drug nomenclature and advertising would enable physicians more easily to become aware of the nature of the drugs they prescribe, of the existence of lower-priced drugs, and of the precise nature of the alternative available in the drug market. The insuring of the adequacy of drug products inspection would render specious all activities devoted to the disparagement of lower-priced drugs.

Uncompromising measures, rigorously enforced, are indicated in order to halt, and in time to remedy, the existing misallocation of resources in excessive selling efforts, duplicative research and product development programs, and exceptionally high profit levels—all of which are characteristics of the industry and much of which is paid for by individuals who can ill afford to do so.

Reprinted from

The Journal of Industrial Economics

Volume XII JULY 1964 No. 3

HENRY STEELE

Patent Restrictions and Price Competition in the Ethical Drugs Industry.

BASIL BLACKWELL

BROAD STREET OXFORD ENGLAND

Price: Thirteen shillings and sixpence net or $1.90. Annual subscription Thirty shillings or $4.20

1971

policy recommendations volunteered by witnesses varied all the way from a request that drug patent protection be extended from the present 17 years to a period of 25 or 30 years,[4] to a suggestion that some personal violence against certain drug makers and their lobbyists might be in order.[5] In part, however, the inconclusiveness of some of the economic aspects of the record is due to the 'trade secret' status afforded certain crucial cost data, and the consequent failure to insist upon the publication of actual production costs for some of the higher-priced patented drugs. To the economist, this is a most unfortunate omission, for such data are otherwise entirely unobtainable.

This paper is directed to the assessment of price competition among ethical drugs, and the extent to which such competition is prevented by the patent privilege. Before undertaking a detailed investigation of price competition in the various drug markets, however, it is desirable to describe the factors influencing supply and demand for ethical drugs.

II. ETHICAL DRUGS: THE MARKET CONTEXT

The most striking features of the ethical drugs market are the extreme inelasticity of demand and the monopolistic restriction of supply, with the element of rivalry (rather than price competition) taking the form of product differentiation through research and development.

As regards the nature of demand, the prescription drugs market is unique. Only a physician can order such drugs, but the patient must make the payment. Chemically identical drugs may be offered for sale under different names at widely varying prices, but the physician has no direct motivation to prescribe the lowest-cost brand, or even to become aware of prices at all. Furthermore, the demand for a drug which is efficacious in a given disease does not simply consist of the total demands of all individuals suffering from that disease at a given time. Instead, the relevant market for such a drug is composed of the total effective demands of all individuals who can be persuaded to consult physicians, and who suffer (or think they suffer) from whatever disorders physicians may be inclined or induced to prescribe the given drug in connection with. 'Ethical drugs' by

[4] Testimony of Dr. Chauncey D. Leake of Ohio State University, ibid., Part 18, p. 10433.

[5] Testimony of Mr. Mike Gorman, executive director of the National Committee Against Mental Illness. In commenting upon a speech by F. C. Brown, president of Schering, and then president of the American Pharmaceutical Manufacturers Association, he concluded: 'In fact, Mr. Brown went much further than Marie Antoinette, and he still has his head and his profits, too', ibid., Part 16, p. 8995.

range of products than is allowed in the original grant. A series of improvement patents may indefinitely prolong the life of an original patent, as has occurred in the case of insulin,[9] and these are usually easy to obtain. But it is also easy to 'patent around' existing chemical patents, and if such efforts result in devising improvements on products made or processes employed by original patent holders, the logical solution is cross-licensing. Cross-licensing negotiations, involving as they do the mutual compromise of patent monopoly positions, supply the motivation for a greater sense of community of interest in price and production policies, and serve to further limit competition.[10] Drug patents may bring about few absolute monopolies, since few patented drugs are without effective substitutes, but the licensing and cross-licensing agreements prevalent among patent holders facilitate a high degree of market control from the supply side. Lack of space prevents effective documentation of patent abuses, but it may be mentioned that of the forty-two most important patented drugs, twenty-four are produced by only a single supplier.[11]

Not all ethical drugs are protected by patents, and since there seem to be no important economies of scale in production for such drugs, it would appear possible for small firms to compete. Hence the larger firms have taken measures (1) to confuse the normal flow of market information in order to prevent physicians from knowing of lower-priced sellers in the market; (2) to prevent the identification of lower-priced equivalents of higher-priced drugs, and (3) to persuade the physician that all lower-priced drugs are of hazardously low quality. These efforts have been quite successful. Small firms with little or no advertising budgets cannot make their presence known in the deluge of major firm drug propaganda.[12] Marketing tactics produce economies of scale in advertising drugs, where none exist in producing them. The second objective is accomplished by making the physician brand name conscious, and by devising and advertising generic names of drugs in such a way as to minimize the use of cheaper equivalents. Generic names are designed to be lengthy and complex; brand names are brief and euphonious. Generic names, by law, must be included in all advertising, but are usually printed in minute type face and located in anomalous places. Brand names are given great prominence and are advertised inten-

[9] Ibid., p. 141.

[10] The currently pending Federal Trade Commission complaint, *FTC* v. *American Cyanamid, et. al.*, charges the respondents with such a conspiracy in connection with the tetracycline patent negotiations, ibid., pp. 145–7. See also the discussion of the prednisone cross-licensing agreements given below.

[11] Ibid., p. 67.

[12] Selling expenses amounted to $3200 *per physician* in 1959, and constituted 24.8 per cent of sales revenue and was 77.3 per cent as large as the cost of goods sold, ibid., p. 31.

sively. If a physician prescribes by generic name, the pharmacist may fill the prescription with any brand of that drug; if he prescribes by brand name, the pharmacist must dispense only the brand called for. Unless he obtains specific permission from the physician, prescribing another brand makes him liable to prosecution by state pharmacy boards under the 'substitution' law.[13]

The third objective is accomplished by a continuing program of disparagement of lower-priced drugs on the part of the major drug firms through their trade associations and salesmen, especially detailmen.[14] In no other industry is it likely that disparagement of lower-priced products has been able to supplant price competition so completely. Food and Drug Administration regulation has been surmounted by also disparaging the adequacy of FDA inspection.

III. PRICE COMPETITION AMONG ETHICAL DRUGS

Very little useful quantitative data exist outside the material published in the *Hearings*, which concerned themselves with four drug categories, antibiotics (about 20 per cent of total ethical drug sales), corticosteroid hormones (10 per cent), tranquilizers (9 per cent) and oral antidiabetic drugs (3 per cent). The corticosteroids market is easily the most representative of the four, since the pattern is not as clear in antibiotics nor as complete in tranquilizers or oral antidiabetics. Consequently, chief emphasis will be given to the corticosteroids market as a paradigm of the effect of patents and other drug marketing institutions on price competition.

A. *Corticosteroid Hormones*

There are seven closely related corticosteroid hormones which are used in the treatment of a number of related disease groups, particularly those with inflammatory symptoms of a chronic nature. All seven are to some extent substitutes for one another in the treatment of many disorders.[15]

Cortisone, the first corticosteroid hormone to be discovered, was found in 1944 and commercially produced in 1948. Its price history

[13] Among pharmacists, 'substitution' formerly signified the dispensing of a chemical substance other than that required by the physician. The National Pharmaceutical Council engaged in a crusade to induce state boards of pharmacy to draft antisubstitution laws defining substitution in terms of brand names, and succeeded in increasing the number of states with such laws from four to forty-four from 1953 to 1959, ibid., p. 235.

[14] Part 18 of the *Hearings* contains much testimony to this effect by physicians and medical educators. See particularly the testimony of Dr. Solomon Garb, *Hearings*, op. cit., Part 18, pp. 10480–5.

[15] A study made in 1959 showed that all seven of these drugs were prescribed in every one of the seven disease groups requiring corticosteroids therapy, *Hearings*, op. cit., Part 19, p. 10825.

reveals a pattern characteristic of a new product in the chemical industry, with initially almost prohibitively high prices falling rapidly to reach a plateau. Merck and Company first developed cortisone, and although unable to obtain a product patent, it managed to keep a pure but transitory monopoly of production until 1952, largely because it alone had processing facilities. In the absence of patent protection, the cost savings from process improvements were passed on to buyers:[16]

July 1949	$200 per gram	Aug. 1950	$50 per gram
Dec. 1949	150 „ „	Oct. 1950	28 „ „
Mar. 1950	135 „ „	Apr. 1951	24 „ „
May 1950	115 „ „	Oct. 1951	20 „ „
July 1950	95 „ „		

But by 1953, Merck's brand, 'Cortone' held only 54.8 per cent of the market. Intense price competition continued through 1954, at which time Merck's prices had dropped to $5.48 per gram. Merck, Schering ('Cortogen'), and Upjohn ('Cortisone') have produced virtually all of the United States supply of cortisone since 1954, and each has charged $5.48 per gram at wholesale since that time.[17] Hydrocortisone prices fell commensurately with cortisone prices until 1956, when all producers quoted identical wholesale prices of $7.99 a gram.[18]

A third corticosteroid hormone, prednisone, was first discovered by Syntex in 1955, but several other firms were working on corticosteroids, and discovered the compound at about the same time. (Merck claims to have produced prednisone shortly after their discovery of hydrocortisone, but failed to realize its potentialities.[19]) Schering, Merck, Upjohn, Pfizer and CIBA all filed patent applications on prednisone, and an interference proceeding was declared by the Patent Office. Syntex later became a party to the interference, basing its claims on a 1951 patent covering a class of products which Syntex claimed included prednisone and prednisolone. The interference has not yet been resolved by the Patent Office, but the parties involved undertook to make interim arrangements among themselves. Apparently Schering could convince Upjohn, Merck and Pfizer that it had the best chance of obtaining the eventual patent. In 1955, Merck, Upjohn and Pfizer separately agreed on cross-licensing arrangements with Schering, each involving the payment

[16] Testimony of J. T. Connor, ibid., Part 14, p. 8030.
[17] Ibid., Part 14, p. 7884.
[18] Ibid., Part 14, p. 7885.
[19] Testimony of J. T. Connor, *Hearings*, op. cit., Part 14, p. 8027.

to Schering of unconditional interim royalties of 3 per cent of sales, for a period of 3 years, in exchange for the right to make prednisone and prednisolone and market it only in finished dosage form. The royalties were 'interim' in the sense that they could be collected by Schering until an interference was declared at the Patent Office, after which time they would simply accrue, but would not be paid to Schering until and unless a patent was issued to Schering. Parke, Davis signed a largely similar agreement in 1957, and CIBA followed in 1958. Such licenses were granted on the basis of patent applications, rather than on patents held.[20] The president of Schering defended the practice on the grounds that this was the only way to insure that all parties involved would be licensed to continue production, no matter who was awarded the patent.[21] This is an admission that failure to pay Schering interim royalties might jeopardize the granting of a license if Schering did obtain the patent, and the size of the interim royalty is a measure of the degree to which the other companies considered Schering's patent claims to be superior to theirs. All of the cross-licensing agreements (except with Upjohn) provided for the sale of prednisone and prednisolone to third parties in finished dosage form only, thus preventing bulk sales of the finished powder to competitors who might tablet the powder and sell it, either generically or under their own brand names, at their own prices. Upjohn, although not bound by such a restriction, has in fact made no sales of bulk powder except to Schering.[22] Schering's president saw nothing unusual in the interim royalty feature. Merck's president disagreed in principle, but conceded that there was probably nothing illegal about it.[23]

From the very beginning, the wholesale prices of both prednisone and prednisolone charged by Schering, Merck, Upjohn, Pfizer, CIBA and Parke, Davis have been identical, at $17.90 per bottle of one hundred 5-milligram tablets.[24] The sixth party to the interference, Syntex, requested a license under Schering's process in 1955. A Schering spokesman is said to have refused because Syntex had a reputation 'for knocking the pants off prices'.[25] Syntex then began to make sales of prednisone bulk powder (i.e. the finished product in bulk powder form) in the United States in late 1956. Schering

[20] Data taken from cross-licensing agreements submitted to the Suncommittee by Schering, ibid., Part 14, pp. 7918–20.
[21] Testimony of F. C. Brown, ibid., Part 14, p. 7928.
[22] *Report*, op. cit., p. 151.
[23] *Hearings*, op. cit., Part 14, p. 8096.
[24] Data from *American Druggist Blue Book*, as reported in the *Hearings*, ibid., Part 14, p. 7884.
[25] Ibid., Part 14, p. 7920. Reference is to a remark allegedly made by Irving Jurow of Schering (who disavows the phrase) and reported in *Fortune*, August 1958.

brought suit for infringement under the process patent, and Syntex retaliated by suing both Schering and Merck for infringement of its 1951 product patent. Schering thought enough of the validity of the Syntex patent to agree in 1958 to pay Syntex a lump sum paid-up royalty in return for a release from all claims, present or future, under the 1951 patent.[26] Merck made a similar agreement with Schering at about the same time.[27] However, the bulk powder price structure for prednisone and prednisolone had already begun to deteriorate by the time such agreements were accomplished. When Syntex began to make its 'uncontrolled' bulk shipments into the United States, Pfizer early in 1957 violated the terms of its cross-licensing agreement with Schering and made bulk sales of its own powder to third parties, meeting Syntex's low price. Merck soon followed suit, but Upjohn continued to sell in bulk only to Schering.[28] This competition in the bulk market steadily forced bulk prices downward, from levels which were initially quite low in relation to the wholesale price of the tableted and bottled powder. The bulk prices charged by Syntex declined from $10.01 per gram in the first quarter of 1957 to $2.36 per gram in the third quarter of 1959. Merck and Pfizer are said to have met Syntex's price; after they entered the market at lower prices they made sales at Syntex's expense, and the records show that sales by Syntex declined, despite the great reduction in its price.

In March 1959 Schering and Syntex made an agreement, under the terms of which Schering agreed, in the event that it was awarded the prednisone patent, to license Syntex (1) to make prednisone and sell it in bulk form to pharmaceutical manufacturers for use as a chemical intermediate in the manufacture of products other than prednisolone, or of those so closely related to prednisone or prednisolone as to constitute their equivalents, and to pay a 6 per cent royalty on such sales; (2) to make and sell prednisone in finished form under its own label at a 6 per cent royalty rate; and (3) to sell prednisone in bulk to licensees of Schering, who in turn would pay royalties directly to Schering. If Syntex were to be awarded the patent, Schering would receive a license to make and sell prednisone at a royalty rate of 3 per cent.[29]

This agreement contrasts with those made by Schering with Merck, Pfizer, Upjohn, CIBA and Parke, Davis. The latter agreements provided for immediate cross-licensing, the payment of a

[26] Testimony of F. C. Brown, ibid., Part 14, pp. 7921–2.
[27] Text of Merck–Schering agreement, ibid., Part 15, pp. 8364–9.
[28] Ibid., Part 14, p. 8095.
[29] Text of agreement submitted by Schering to Subcommittee, ibid., Part 15, pp. 8878–9.

3 per cent interim royalty (5 per cent for Parke, Davis), such interim royalties to be in effect for 3 years (the earlier agreements with Merck, Pfizer and Upjohn stipulated that in the event of an interference proceeding,[30] interim royalty payments would be suspended and allowed to accrue until a patent was actually issued; the CIBA and Parke, Davis agreements called for interim royalties for 3 years after the date of the agreement itself); and, finally, all of the agreements (except with Upjohn) restricted sales to finished dosage forms only.[31]

The purpose of restricting sales to finished forms only is clearly to prevent any shipment of bulk powder being made available to non-licensed competitors (chiefly the smaller firms who take the bulk chemical, tablet and bottle it, and sell it under their own brand name, or more frequently, under the generic name) who might make use of such an 'uncontrolled' supply to introduce price competition into the market. It is by no means unusual in the drug trade to find licensing agreements predicated upon the basis of patent applications rather than issued patents, but it is somewhat unusual to have interim royalties paid on the basis of applications alone. Counsel for the companies pointed out that there was nothing illegal about interim royalties, since one firm might very well pay another a royalty in order to use a secret process for which no patent was sought. In such circumstances, the consideration for the royalty would be the disclosure of an otherwise unknown process. Schering's agreements with Pfizer, Upjohn, Merck and perhaps some of the others, could not be so felicitously construed, since all firms had been doing intensive research in the field and were roughly on the same footing.

It was generally agreed, however, by Schering and its licensees, that it would be impossible to enforce the restrictions on the form in which the good is sold, until a valid patent had been issued.[32] Nevertheless, a pattern of action which can be described as voluntary compliance was observed until Syntex began making its shipments of bulk prednisone, at which time Pfizer and Merck violated the letter of their agreements with Schering by following suit. The initial effort at voluntary compliance may conceivably have been a gesture of goodwill toward Schering, or it may have been a collectively acceptable way of dealing with the problems posed by highly profitable prices (which could bear a great deal of shading) coupled

[30] The prednisone and prednisolone interferences were declared in December 1956 and December 1957 respectively, ibid., Part 14, p. 8093.

[31] Licensing agreements submitted to the Subcommittee, ibid., Part 15, pp. 8361–4.

[32] Testimony of F. C. Brown, ibid., Part 14, pp. 7926–8; testimony of J. T. Connor, Part 14, p. 8094.

with the probable excess capacity on a product which as yet had no patent protection.[33] When Pfizer and Merck violated their agreements by making bulk sales, Schering took no action, and its president indicated in the hearings that he realized that this part of the agreement was not enforceable.[34]

Price competition ensued. Small companies could obtain the bulk powder from Syntex, Merck, Pfizer or from a small domestic producer, Formet, and package their own finished dosage forms with no restrictions until such date as a patent might be issued. If a patent were to be issued to any firm other than Syntex, these smaller producers could expect to be eliminated from the market immediately.[35] Until such time, however, small firms were free to cut prices as far below those of the major firms as they could. The lowest wholesale price reported in the hearings ($1.75 per hundred 5-milligram tablets) was less than one-tenth of the $17.90 charged by Merck, Schering and Upjohn.

Given perfect market information, differences of 90 (or of 1000) per cent in the prices of identical products would be impossible. The larger firms found it much more congenial to disparage the products of lower-priced sellers than to quote lower prices to meet their competition. Nor was any price reduction necessary in the retail druggist prescription market, where well over 90 per cent of the sales were made by the three large firms. It may be appropriate to inquire, in the perhaps extreme case of prednisone, what substance there may be to charges that lower-priced producers sell inferior prednisone.

Prednisone could be purchased under the generic name in late 1959 from a number of smaller firms, for prices ranging from $1.75 to $12.00 per hundred 5-milligram tablets at wholesale. It could also be purchased from Schering as 'Meticorten', from Upjohn as 'Deltasone', from Merck as 'Deltra', or from Parke, Davis as 'Paracort'

[33] Excess capacity is very likely to have existed. Dr. Philip Berke of Formet Laboratories testified that an investment of only four or five million dollars would easily be sufficient to supply the entire world demand for prednisone and prednisolone, ibid., Part 14, p. 8056.

[34] Ibid., Part 14, p. 7928.

[35] It has not been the practice of the major drug firms to license smaller companies. Blackman of Premo, a small firm selling prednisone and other drugs at low prices under generic names, stated that he could produce the newer, and as yet unpatented, corticosteroid dexamethasone in the interval before a patent was issued, but that after its issuance his investment in production facilities would be worthless. He declared: 'With as much assurance as any human being can muster, I feel we would not get a license', ibid., Part 14, p. 8231. Dr. Berke of Formet Laboratories, which is apparently the only small firm in the United States making prednisone and prednisolone in bulk form, was of the same opinion, and indicated that he would apply for a dexamethasone license 'tongue in cheek'. He further indicated that it was cheaper to produce prednisone and prednisolone than dexamethasone; hence he could afford to produce the former while patent proceedings were still pending, but not the latter, ibid., Part 14, pp. 8057–8.

at $17.90. Claims that brand name differences mean quality differences must be dismissed. Prednisone is listed in the *United States Pharmacopoeia*, which specifies standards of purity, potency and identity which must be met by all makers, and there is no therapeutic gain in producing drugs to a higher purity exceeding them.[36] They are enforced by Food and Drug Administration inspection. No producer of prednisone, large or small, had ever been accused of irregularities in connection with prednisone products as of the date of the hearings. For prednisone, all sellers must stay within the potency tolerance, between 90 and 110 per cent of the level claimed on the label, whether their products are branded or are sold generically. Since the cost of the active ingredient is typically about one-third of total factory cost, there can be no appreciable savings in aiming at the lower limit. Furthermore, the firm that sells the product may not manufacture the bulk powder, or even do the tableting. Parke, Davis, in 1958, made none of its own bulk powder, while the smaller companies bought their bulk powder largely from Pfizer and Merck (who are reported to have met Syntex's lower prices) as well as from Syntex and Formet. Further evidence of the reliability of at least some of the lower-priced sellers is given by the fact that both Premo and Chase were allowed to make sales to the Military Medical Supply Agency, which makes thorough inspections of the facilities of all low bidders on government drug purchase contracts.[37]

No brand name preference is given for purchases by government agencies or by hospital formularies. Here, small firms may compete with large firms in terms of price, and large firms, especially on government contract bids, may have to cut their prices drastically, and in such a way that something may be learned of their costs and pricing policies. Sales were negotiated by the Veterans' Administration with Merck and Schering in February 1958 at a price of $136 per bottle of one thousand 5-milligram tablets of prednisone. In March 1958 competitive bidding was instituted; Merck cut its bid to $95, Schering dropped to $68, but a small firm, Panray, obtained the contract with a low bid of $38.50, or 28.3 per cent of the price negotiated the month before. Competitive bids submitted to the Military Medical Supply Agency at about the same time showed Merck bidding $75, Schering $79.74, and another small firm, Chase, obtaining the contract with a low bid of $41.50. In April 1958 Merck and Schering entered their previous bids, and Premo was the low bidder at $38.40. In December of the same year, Schering cut its

[36] Testimony of Dr. Walter Modell of Cornell University, ibid., Part 21, p. 11610.
[37] Testimony of Captain H. R. Fahlbusch of the Military Medical Supply Agency, ibid., Part 21, p. 11547.

price to $23.62 and obtained the contract, underbidding smaller firms. At this time, Schering was charging $170 per thousand at wholesale to druggists, implying a discount of 86.7 per cent to the government.[38] In February 1959 Premo was again the low bidder, with a price of $20.98. At that time, Premo was selling at $26 per thousand to the druggist, with the comparative discount to the government at 19.3 per cent.[39]

It is interesting that quoted wholesale prices and bids tend to vary inversely with the size of firm. The three large firms charged identical prices of $17.90 per hundred 5-milligram tablets. Prices of the smaller companies named varied from Lannett's $12.00 to Premo's $2.35. A Washington chain drugstore dealer reported to the Subcommittee that he had been able to purchase the same goods at $1.75 from an unidentified 'first line' company, and the 1959 *Drug Topics Red Book* listed twenty firms selling at prices below $3.00.[40] Such facts lead one to the question: what is the relationship between size of firm and cost of production? The presidents of Merck, Schering and Upjohn testified that the sales they made to the government were either not profitable, or did not yield a sufficient profit to cover the cost of an adequate research program. The smaller firms testified that they definitely made a profit at their prices. Syntex, however, spent about half again as high a percentage of sales revenue upon research as did the larger firms, and Formet spent relatively nearly twice as much.[41] A comparison of selling costs is more instructive. Premo (the only small firm for which sales costs are available) devotes 2.0 per cent of its sales dollar to selling activities while the weighted average for the three large firms was 24.3 per cent.[42] There is some reason to suspect that Premo might spend more of its revenues on selling efforts than do other small firms. Premo once attempted to increase its sales force, hire detailmen and compete with the large firms, using their own methods. In 1955 the company's sales outlays increased to almost 7 per cent of sales, but their campaign failed because of 'the tremendous increase in the advertising dollars spent by our large competitors, to the extent that our efforts appeared, in the market place, as a mere spark in a vast conflagration'.[43] Advertising outlays were cut, but may yet remain above the levels for small firms which never attempted large-scale selling campaigns.

[38] Data submitted by the Military Medical Supply Agency to the Subcommittee, ibid., Part 21, pp. 11551–3.

[39] Data from Premo 1959 Hospital Price Catalog, as reported in the *Hearings*, ibid., Part 15, p. 8709.

[40] *Report*, op. cit., p. 17.

[41] *Hearings*, op. cit., Part 14, pp. 8064, 8301.

[42] *Report*, op. cit., p. 31.

[43] *Hearings*, op. cit. Part 14, p. 8215.

Production costs for prednisone are only a small fraction of the wholesale price. Rather than make public the actual production costs of individual firms, the Senate Subcommittee staff estimated the cost of a bottle of one thousand 5-milligram tablets of prednisone at $13.61.[44] This cost estimate was attacked by industry witnesses as being too low, as not including selling and research costs, etc., etc., despite its being made unmistakably clear in the record that the cost estimate referred only to production costs. Actually, this cost estimate is quite generous. It is based on a bulk powder cost obtained by using the prices charged by bulk producers for sales of bulk powder to other firms, and hence includes not only their actual direct and allocated costs, but also some margin for profit. To determine the cost of tableting and bottling, firms which perform such functions on a contract basis were contacted, and the median cost quotation (not the lowest) was employed. The actual production cost of McKesson and Robbins is only $8.99 for such a bottle.[45] This cost is only 5.07 per cent of the equivalent wholesale price of $179 per thousand, and only 3.17 per cent of the retail price of $283.33.[46]

Price competition departs and then reappears on the scene with regard to the three newer corticosteroids. When Upjohn introduced methylprednisolone in 1957 ('Medrol'), it first had the market to itself, and only later licensed Schering, which had no chance to market its brand until 1959. Being 25 per cent more potent than prednisolone (a standard dosage form is the 4-milligram tablet as compared with the 5-milligram prednisolone tablet), its price per milligram was increased by exactly 25 per cent, so that the price per tablet is the same, and the introduction of the more potent product did not give rise to any price competition.[47] As a new drug, however,

[44] *Report*, op. cit., p. 34.

[45] Testimony of Dr. Herman Nolen, president of McKesson and Robbins, in *Hearings Before the Antitrust Subcommittee of the Committee on the Judiciary, House of Representatives, Eighty-seventh Congress, Second Session*, Washington, Government Printing Office, 1962, p. 15.

[46] While the present paper is not concerned directly with the retailer's markup of 66⅔ per cent, it certainly helps to contribute to the magnitude of the final price to the consumer. It is well known that retail druggists, like retailers everywhere, dislike price competition in practice, and favor high markups—the role of the National Association of Retail Druggists as one of the prime movers behind the 'Fair Trade' laws is notorious—and the complaints that high markups have failed to produce satisfactory profits for retail druggists is simply a consequence of the freedom of entry in the field and the overcapacity, low turnover rates and higher average unit costs which results from new entry in response to the high unit margins. It is pertinent to call attention to this point in view of the possibility that pressure may be exerted by organized druggists on the manufacturers to discourage any tendencies toward price cutting, or at least that the drug manufacturers may be aware of such distributor sentiment when contemplating possible price cuts. It must be admitted, however, that no shred of evidence on this point was offered (or sought) during the whole of the hearings.

[47] Data from *American Druggist Blue Book, 1959–1960*, as reported in the *Hearings*, op. cit., Part 14, p. 8324.

methylprednisolone began to take over a notable share of the market, and by 1958 'Medrol' had 18.1 per cent of the new prescription market, while the leader in 1956, 'Meticorten', had fallen from 53.9 per cent to 30.7 per cent.[48]

In 1957 Lederle introduced triamcinolone, a corticosteroid identical in potency to methylprednisolone, and licensed it to Squibb. Both firms set prices at a level identical with that of prednisone, prednisolone and methylprednisolone. Lederle's brand, 'Aristocort', had 23.8 per cent of the new prescription market in 1958, while Squibb's brand, 'Kenacort', had 8.9 per cent.[49] Eighty-five per cent of the bulk powder, however, was made by Squibb, and only 15 per cent was made by Lederle. As a consequence of Lederle's prior and intensively advertised appearance in the market, it would appear that physicians' brand preferences crystallized in favor of Lederle's product so predominantly that Squibb had no outlet for much of its production except to sell it in bulk to Lederle.

In 1958 Merck and Schering introduced dexamethasone, a corticosteroid with a potency about six and two-thirds as great as prednisone, the standard tablet containing .75 milligrams instead of 5 milligrams. For the first time since the prices of cortisone and hydrocortisone reached a plateau, price competition was introduced into the market when Merck priced its 'Decadron' at retail at 27 cents per tablet, 10 per cent below the prices of the other four competitive corticosteroids. Schering later marketed its brand, 'Deronil', at the same price as Merck. Merck, intending to regain its earlier predominant place in the market, gave its dealers additional quantity discounts amounting to as much as 10 per cent.[50] A patent interference was declared between Merck and Schering. Merck subsequently licensed CIBA and Organon of Holland. CIBA then began to market dexamethasone as 'Gammacorten' in 1958. While dexamethasone has yet to be patented, no small firms seem to be producing this compound, because of the relatively high costs involved. In fact, it may be the case that the profit margins to the producers are considerably lower than in the case of the other corticosteroids. Sales of bulk powder by Merck to CIBA were made in 1958 at $65 per gram. If the bottling and tableting charges used previously are added to the bulk powder cost, a cost of $72.69 per gram is obtained, which amounts to 33.8 per cent of the wholesale price of $214 per gram.[51] (The comparable figure for prednisone is 8.7 per cent.) This is only slightly more than the average ratio of

[48] Data from the market research staff of Merck and Company, ibid., Part 14, p. 8028.
[49] Ibid., Part 14, p. 8028.
[50] Testimony of J. T. Connor, ibid., Part 14, p. 8031.
[51] Ibid., Part 14, p. 8324.

32.1 per cent for cost of goods sold to total sales for the twenty-two largest firms in the industry, as cited above. Dexamethasone therefore combines a lower price with much higher costs—a phenomenon impossible in pure competition.

It is inappropriate to attempt to trace in corresponding detail the course of price competition, product by product, and in the light of existing patent licensing agreements, in the other drug group markets. They all reveal the same general features, hence only a sketch will be given of prices, the relation of prices to computed costs and competitive bids, and of patents, in the other three markets.

B. *Oral Antidiabetic Drugs*

The oral antidiabetics market may readily be summarized in this regard. Each of the three products is either patented or produced under patent application arrangements. Upjohn set the price of 'Orinase' in such a way as to be almost exactly equal to the price of insulin on a per dose basis, at 13.9 cents per 500-milligram tablet, comparable to 14.0 cents for 10 units of insulin.[52] This price was not based on cost, for the cost of insulin is no doubt considerably greater than the cost of 'Orinase', or tolbutamide, since animal pancreas is undoubtedly more expensive to obtain and process than the industrial chemicals required to produce tolbutamide.[53] Yet the greater convenience of administration of an oral drug over an injectable drug would probably support a higher price than the insulin equivalent. The price was simply set equal to insulin's price, apparently in deference to the industry's tradition of pricing new products at (or very near) the price of existing drugs used to treat the same disorder. (The price set is actually below the total cost of insulin therapy, including syringe and needle.) The price of insulin itself had not been changed since 1947. Had the practice of medicine not changed, we might still be paying second-century A.D. prices for the remedies of Galen.

Pfizer priced its drug, chlorpropamide ('Diabinese') at 15 cents per 250-milligram tablet, but since Pfizer's drug is somewhat more potent than tolbutamide, the daily cost may run somewhat less. The third drug, phenformin, is made by one of the smaller firms, U.S. Vitamin. Phenformin is still more potent than chlorpropamide, a 50-milligram tablet costing 12 cents. The range of possible daily cost (minimum to maximum recommended dosage) is from 14 to 56 cents for tolbutamide, from 7 to 30 cents for chlorpropamide, and from 12 to 55 cents for phenformin.[54] The drug buyer seems to have

[52] Testimony of Dr. Upjohn, ibid., Part 20, pp. 11037–9.
[53] *Report*, op. cit., p. 99.
[54] *Hearings*, op. cit., Part 20, p. 11120.

real price alternatives, for once, but because of the priority in appearance of the higher priced drugs, and their intensive advertising, sales seem to be directly related to the height of daily average cost.

The Subcommittee staff computed the estimated production cost for one thousand 500-milligram tolbutamide tablets (twenty bottles of fifty each) on the basis of bulk powder transactions plus contract tableting and bottling charges, as in the case of prednisone. The production cost estimate is $6.86; the royalty (7½ per cent of the selling price) is $6.24. The wholesale price of twenty bottles is $83.40 and the retail price is $139.00 Upjohn's production cost is therefore 8 per cent of the wholesale price; the production cost plus the royalty brings Upjohn's total estimated cost of production and royalties to 15.5 per cent of the wholesale price and to 10.6 per cent of the retail price.[55] Production cost estimates for the other two drugs are not available. The only bid on a Military Medical Supply Agency contract that is available for any such drug is Upjohn's successful bid of $4.00 for 18,432 bottles of fifty 500-milligram tablets. Since the wholesale cost for one bottle is $4.17, it can be seen that in the absence of lower-priced competition from smaller firms, a large order from the government, with its attendant cost savings, could be obtained at only a 4 per cent quantity discount.[56]

C. *Tranquilizers*

Price competition among tranquilizers varies with the type of tranquilizer. There are two types, the 'potent' type used chiefly in mental institutions, and the 'mild' type used for non-hospitalized patients troubled with anxiety.[57] In the potent group, the first drugs to be introduced were chlorpromazine and reserpine in 1954. Both drugs were priced at wholesale at $3.03 per bottle of fifty tablets. Later during the year CIBA introduced its brand of reserpine at a price for fifty tablets 25 per cent below its existing competitors for that drug (Squibb and Riker), $2.25. When Smith, Kline and French introduced prochlorperazine in 1956, they priced it at $3.03 per fifty tablets, the same price they charged for their chlorpromazine. When Wyeth introduced promazine hydrochloride in 1957, it did so at a price of $3.00 for a bottle of fifty tablets.[58] In an interesting

[55] *Report*, op. cit., p. 20.

[56] *Hearings*, op. cit., Part 21, p. 11551.

[57] As is true in most drug categories, much substitution among drugs is possible in most disease groups treated by tranquilizers. A 1959 study by the Pharmaceutical Manufacturers Association shows that all seven of the major tranquilizers (both mild and potent) are prescribed in all five major disease groups requiring such therapy, ibid., Part 19, p. 10827.

[58] Ibid., Part 16, p. 8887.

move in 1958, Wyeth changed the wholesale price of its drug promethazine hydrochloride, which was originally an antihistamine with drowsiness-inducing properties, from $4.45 per hundred 25-milligram tablets to $5.94, thus making it competitive with its recent tranquilizer promazine hydrochloride.[59] An antihistamine was thus made a tranquilizer in response to changing market demand by changing the advertising to stress what were previously considered side effects as the desirable end result, and by increasing the price in recognition of the increased marginal revenue possible in the more lucrative market for tranquilizers. Later potent tranquilizers were priced at about $3.00 for the equivalent dosage. Apparently reserpine was considered to be a sufficiently distinct product that CIBA's 25 per cent lower price did not call for price reductions for other potent tranquilizers. CIBA, perhaps uncertain as to the validity of its patent (reserpine is the name given to the active substance isolated from the root of *Rauwolfia Serpentina*, a species of tropical plant, the medicinal properties of which had been known in India for centuries), issued licenses to nine other firms in the United States, eight of which were allowed to sell in bulk powder form. These firms made sales to smaller firms, which in turn proceeded to cut prices down to as much as 90 per cent below CIBA's levels. CIBA continued to advertise its brand of reserpine and its price held up; other major firms dropped their advertising of this product, by and large, and their prices were cut to about $1.91 per fifty tablets at wholesale.[60] No other potent tranquilizer has ever experienced a cut in price.

Carter was first in the mild tranquilizer market with meprobamate, which it has sold since 1955 at $2.60 per bottle of fifty 400-milligram capsules at wholesale. Carter licensed Wyeth to sell meprobamate under Wyeth's brand name, and Wyeth charged the same $2.60, although its royalty payment to Carter would be 13 cents, and its raw material cost (purchased from Carter) would be 25 cents greater.[61] Carter licensed its very profitable product only because it did not have the selling facilities which larger drug firms such as Wyeth had. Little competition to meprobamate has developed. Phenaglycodol is sold by Lilly at the same price to the drug buyer as meprobamate.

The computed cost for reserpine production is $2.48 per bottle of one thousand 25-milligram tablets.[62] CIBA's wholesale price for this quantity is $39.50; other firms quote prices as low as Panray's $2.65. For CIBA, estimated production cost is 6.3 per cent of its

[59] *Hearings, op. cit.*, Part 16, p. 9274.
[60] Ibid., Part 16, pp. 9387–90.
[61] *Report*, op. cit., p. 18.
[62] *Hearings*, op. cit., Part 16, p. 9436.

price; for Panray, 93.6 per cent. It must be the case, however, that the computed cost figure is much too high, for in competitive bids to the Military Medical Supply Agency Panray has won bids at as low as $0.51 per thousand tablets on a 4080-bottle order, and CIBA has failed to get contracts on bids as low as $0.52—1.3 per cent of the wholesale price. When asked how CIBA could afford to bid successfully at $0.60 in February 1959 Mr. T. F. D. Haines, president of American CIBA, testified: '. . . we didn't anything like recover our out-of-pocket costs . . . it was perhaps a mistake that we did that'.[63] This makes it all the harder to explain CIBA's later and lower unsuccessful bids. (Raw materials costs for one thousand 25-milligram tablets should be about 30 cents.) The president of McKesson and Robbins later volunteered his firm's production costs for reserpine: $0.63 per thousand tablets, or 1.6 per cent of the wholesale price.[64]

The computed production cost for meprobamate to Carter is $7.32 for twenty bottles of fifty 400-milligram tablets, on the basis of a 20,000-bottle order. For Wyeth, the cost would be $12.40 plus $2.60 royalty. (Carter agreed that the Subcommittee staff's cost estimate was correct for its own production.) Both Carter and Wyeth obtain $52.00 at wholesale for twenty bottles; a retail drug buyer would pay $108.40 for this amount. Carter's production cost is 14.1 per cent of the wholesale price, Wyeth's is 24.0 per cent, and with the royalty to Carter, 28.0 per cent. (Carter's production cost is only 6.8 per cent of the retail price.)[65] Bids to the Military Medical Supply Agency from Carter and Wyeth ranged from $22.50 for five hundred 400-milligram tablets in February 1958 to $19.845 in February 1960, or from 90 to 79.4 per cent of the wholesale price for this amount. In the absence of any domestic competition, the Military Medical Supply Agency purchased abroad in June 1960, obtaining a low bid of $3.95 from a Danish firm, a bid only 29 cents above Carter's production costs, and about 80 per cent below the bids of Wyeth and Carter.[66] Danish costs may be below American costs, but Carter could at least have met this bid had it seen fit to do so.

D. *Antibiotics*

Price competition in the antibiotics market is intense in unpatented drugs, such as the earlier penicillins, and streptomycin, the patent

[63] Ibid., Part 16, p. 9430.
[64] *House of Representatives Hearings*, op. cit., p. 16.
[65] *Report*, op. cit., p. 18.
[66] Testimony of Rear Admiral W. L. Knickerbocker of the Military Medical Supply Agency, *Hearings*, op. cit., Part 24, pp. 13784–5, 13794.

for which was licensed by Merck to several other firms. Price competition is, however, considerably less in evidence for the later patented antibiotics.

Table I shows how progressively more efficient production methods and continual competitive pressures consequent upon free

TABLE I

PENICILLIN, STREPTOMYCIN AND BROAD SPECTRUM ANTIBIOTICS PRICES, 1945–60

Date (year; month for price changes)	*Penicillin (10 million units, bulk)*	*Streptomycin (10 grams, bulk)*	*Broad spectrum antibiotics (16 250-milligram capsules)*				
			Chlortetracycline	*Chloramphenicol*	*Oxytetracycline*	*Tetracycline (5 brands)*	*Demethylchlortetracycline (150 mg.)*
	$	$	$	$	$	$	$
1945 May	60.00						
1946 May	36.00	160.00					
1947 January	21.00	45.00					
1948 July	13.00	10.24					
1948 December	9.50	8.19	15.00				
1949 February	9.50	8.19	10.00				
1949 March	7.50	5.00	10.00	10.00			
1950 February	4.75	3.15	8.00	8.00			
1950 April	4.75	3.15	8.00	8.00	8.40		
1950 May	4.75	3.15	6.00	6.00	8.40		
1950 November	4.75	3.15	6.00	6.00	6.00		
1951 September	3.75	3.15	6.00	6.00	5.10		
1951 November	2.50	3.15	5.10	5.10	5.10		
1952	1.15	3.24	5.10	5.10	5.10		
1953	.95	1.70	5.10	5.10	5.10	5.10	
1954	.75	1.70	5.10	5.10	5.10	5.10	
1955	.44	.90	5.10	5.10	5.10	5.10	
1956	.52	.75	5.10	5.10	5.10	5.10	
1957	.79	.88	5.10	5.10	5.10	5.10	
1958	.70	.88	5.10	5.10	5.10	5.10	
1959	.28	.38	5.10	5.10	5.10	5.10	5.10
1960	.21	.36	5.10	5.10	5.10	5.10	5.10

Source: Federal Trade Commission, *Economic Report on Antibiotics Manufacture*, Washington, 1958, Section II, pp. 164, 179, 180, 192; *Senate Hearings on Administered Prices in the Drug Industry*, Part 24, p. 13664.

entry made possible the precipitous decreases in the prices of the extremely important drugs penicillin and streptomycin. 'Broad spectrum' antibiotics prices have acted very differently. Lederle quoted an initial price of $15.00 for sixteen 250-milligram capsules of chlortetracycline, possibly overestimating the inelasticity of demand. Within 2 months, Lederle reduced its price to $10.00, at which price Parke, Davis entered the market in March 1949. On February 1st, 1950, both sellers cut prices to $8.00. In April 1950 a new rival, Pfizer, entered the market with oxytetracycline, selling

at $8.40, 5 per cent above its competitors. In May 1950 Lederle and Parke, Davis reduced prices to $6.00, but Pfizer did not meet the cut until November. On September 27th, 1951, Pfizer lowered prices to $5.10, and by November 1st Lederle and Parke, Davis had met the price cut. Tetracycline and demethylchlortetracycline came later, and were also priced at $5.10. Prices were then stable and rigid for almost 9 years, when Pfizer in August 1960 again assumed the role of price leader by announcing a 15 per cent 'trade discount' adjustment which eventually became a 15 per cent price cut, to $4.34.[67] (This price cut was initiated only 1 month before the Senate Hearings on antibiotics were scheduled to take place.) From 1948 to 1949, broad spectrum prices dropped by 33 per cent; from 1949 to 1951, by 49 per cent, and from 1951 to August 1960 not at all. It is reasonable to suppose that since 1951 the costs of producing broad spectrum antibiotics have declined by approximately as much as the cost of producing penicillin, the production methods employed being largely identical.[68] A comparison of the decline of 90 per cent since November 1951 in penicillin prices, with the 15 per cent decline for broad spectrum antibiotics measures the effect of restriction of entry in broad spectrum antibiotics.

The Senate Subcommittee staff made production cost estimates for tetracycline and chloramphenicol, and in 1961, during the House of Representatives hearings, Senator Kefauver made public Bristol's actual production costs. Bristol incurred production costs of $1.67 per bottle of one hundred 250-milligram capsules (in comparison with $2.88 as estimated by the Senate Subcommittee staff in the Senate hearings). The price of such a bottle to the druggist is $30.60; to the consumer, $51.00. Bristol's production cost is 5.5 per cent of the price to the druggist; the addition of royalties to Lederle and Pfizer of $2.15 paid per bottle brings the cost up to 12.5 per cent. (Bristol actually received an average of $25.27 per bottle for such sales in 1958; hence its production costs and royalties totalled 15.1 per cent of its average price received.) Production costs are about 28.7 per cent of the wholesale price for Upjohn, which purchases bulk tetracycline from Bristol; its total production costs and royalties are $9.30, or 30.4 per cent of the price to the druggist. For Pfizer,

[67] Mr. Duncan, the manager of Lederle, construed Pfizer's action as vigorous and aggressive price competition. He testified that Pfizer had cut its prices on a Saturday in order to 'steal a march on the industry'. This explanation apparently surprised the Subcommittee chairman:

Senator Kefauver: 'You mean after ten years they would suddenly steal a march on you?'

Mr. Duncan: 'Yes, senator. These things happen very rapidly. If you can get any kind of advantage on your competitors, you try to do so', ibid., Part 24, p. 13728.

[68] *Report*, op. cit., p. 82. Chloramphenicol costs have probably declined by still more since Parke, Davis discovered a way of making this drug synthetically.

which holds the patent and pays no product royalties, and which reportedly has a lower cost production process, the total costs of production might perhaps be as low as $1.50. If Pfizer pays the same royalty as Bristol for the use of Lederle's process, its production costs plus royalties would be about $2.75. For Upjohn, Bristol and Pfizer, the ratio of production cost plus royalties to the price to the druggist might be about 30 per cent, 15 per cent and 9 per cent.

The gap between estimated production cost and price for chloramphenicol is predictably handsome. One bulk powder transaction did take place, between Parke, Davis and an Italian firm, Farmitalia, at $30.00 per kilogram. Using this price for the raw material, and Upjohn's actual costs for capsuling, finishing, and packaging, a cost of $1.52 for a bottle of one hundred 250-milligram capsules is obtained. This is 5.0 per cent of the price to the druggist of $30.60, and 3.0 per cent of the retail price to the consumer.[69]

The bids received by the Military Supply Agency reflected the status of market competition. For penicillin, competitive bidding was the rule, the price for 250,000-unit potassium penicillin G tablets falling from $1.29 per hundred in April 1957 to $0.65 in June 1960. As in the case of prednisone, small firms selling penicillin regularly charge much lower prices than large ones. For this particular product, the smallest firm (sales of less than $100,000) had the second lowest price ($3.30); the second smallest firm (sales of less than $1,000,000) had the lowest price; and three of the largest firms (sales of $150 to $250 million) had the highest price ($12.00).[70] For broad spectrum antibiotics, the picture is entirely different, as can be seen from Table II.

For this entire period the price to the druggist on all these drugs was $30.60, and the price to the consumer, $51.00. Lederle and Parke, Davis began the period of rigid prices to the trade by quoting a 50 per cent discount to the Armed Services Medical Procurement Agency (the predecessor to the Military Medical Supply Agency); Pfizer allowed about a 51 per cent discount off the druggist's price. Apparently the three firms recognized the high degree of substitutability among their drugs, for prices fell for each seller with

[69] Harry Loynd, president of Parke, Davis, objected at length that the Farmitalia price was much lower than American production costs. In order properly to appraise the reliability of Loynd's assertions, it is advisable to read the whole of his quite remarkable testimony; nevertheless, a consideration of the import agreement is instructive. Farmitalia agreed to supply Parke, Davis with up to 30,000 kilograms of chloramphenicol during 1960 at $30 per kilogram. Through July 11th, however, only 6000 kilograms had been imported. If this price were such a bargain, why did Parke, Davis not take greater advantage of it? The suspicion is that Parke, Davis might be able to make it for less in Detroit. However, even if the price were twice as high, it would only increase the cost per 100-capsule bottle to $2.31, or 7.5 per cent of the sales price, *Hearings*, op. cit., Part 24, pp. 13971–8.

[70] *Report*, op. cit., p. 84.

regard to its previous price until April 1956 when Parke, Davis repeated its previous bid of $12.50. At that time Lederle's price was $11.00 on the first sale of tetracycline hydrochloride, pricing it at the same level as its chlortetracycline. Lederle was low bidder, and was much surprised to find that Pfizer and Squibb had both bid $19.58, and Bristol, $18.97.[71] For the next 2 years Lederle always bid

TABLE II

BROAD SPECTRUM ANTIBIOTIC PRICES TO THE ARMED SERVICES MEDICAL PROCUREMENT AGENCY AND TO THE MILITARY MEDICAL SUPPLY AGENCY, NOVEMBER 1951–JUNE 1960
(100 capsules, 250 milligrams)

Date	*Chlortetracycline*	*Chloramphenicol*	*Oxytetracycline*	*Tetracycline-hydrochloride*
(ASMPA)				
1951 November	15.30	15.30	15.00	
1952 March	15.00			
1953 June		12.50		
1954 January			12.84	
1954 May	12.00			
1955 March			11.47	
1956 April	11.00	12.50	10.97	
1956 October				11.00
(MMSA)				
1957 February	11.00		10.97	
1958 February		12.50		17.24
1958 March				19.19
1958 April		11.25		
1958 June	11.00	11.25	10.75	17.24
1958 November		11.25		17.15
1959 June		11.25		14.36
1959 August				17.15*
1959 December				8.15
1960 May			10.11	6.16
1960 June				5.62

* Low bid—not accepted.

Source: Federal Trade Commission, *Economic Report on Antibiotics Manufacture*, Washington, 1958, p. 194 (1951–56 data); *Senate Hearings on Administered Prices in the Drug Industry*, Part 24, pp. 13779–82; 13791–2 (1957–60 data).

$19.58. On the second bid for tetracycline hydrochloride, Pfizer met Lederle's previous bid of $11.00, but Lederle and Squibb were both at $19.58. Pfizer, like Lederle, then learned that it was not necessary to make 'ridiculously low' bids. Its prices on the next two bids went up to $17.24 and $19.19. In April 1958 Parke, Davis cut its chloramphenicol price by 10 per cent, to $11.25, after a conference with the Agency purchasing officer on the subject of its

[71] Duncan of Lederle conceded that he had made a mistake in simply pricing tetracycline hydrochloride at the same level as chlortetracycline. He characterized the pressures of competitive bidding in terms of carelessness: 'In other words, one gets a little sloppy in bidding for this kind of business. You sometimes simply bid a ridiculously low figure', ibid., Part 24, p. 13690.

rigid price.[72] In June 1958 the tetracycline hydrochloride price dropped back to $17.24, where it stayed for three consecutive awards. In November Squibb cut its bid to $17.15, to undercut the others, all of which bid $17.24. In June 1959 Lederle bid $17.24, Bristol and Squibb bid $17.15, but Pfizer was low bidder at $14.36. Pfizer's bid is surprisingly low,[73] and it encouraged Military Medical Supply Agency officials to believe that prices had at last fallen. Such expectations were disappointed, for in August 1959 three bids were received at $17.15. Negotiations failed to shade the price, and the contract was cancelled because of the Agency's judgment that the price was unreasonably high. Foreign bids were received, and an Italian firm was awarded three successive contracts at $8.15, $6.16 and $5.62 in 1959 and 1960. Rather than lower its price to compete with foreign bidders, Pfizer first protested the Italian procurements (the Comptroller General rejected Pfizer's contentions as being without merit) and then attempted unsuccessfully to buy out the Italian firm.[74] Had Pfizer chosen to engage in price competition, it could in all probability have undercut the Italian firm on all three bids.

The case of tetracycline hydrochloride prices is rather puzzling in that the product is a close substitute for three other drugs sold at lower prices to the government. Perhaps the industry wished to receive higher prices for the new drug, and a certain amount of trial and error bidding was necessary to facilitate the learning process. It is interesting to note that some sort of division of the total tetracyclines market in total dollar sales was accomplished—even at varying prices—during the period October 1956 through October 1959, or from the first award until the date of the first foreign procurement. During this period Pfizer (the patent holder) received 46.6 per cent of all contracts in dollar terms; Lederle received 17.8 per cent; Bristol, 17.6 per cent; Squibb, 17.5 per cent; and Upjohn, only 0.5 per cent. This may indicate nothing more than the varying

[72] Ibid., Part 24, p. 13782.

[73] Senator Kefauver expressed surprise that Lederle would continue to bid $17.24 after the last sale had been awarded at $17.15. Mr. Duncan replied that he was against quoting lower bids because '. . . we had been doing that, and all that happened was that the price just went lower time after time until it got down to an uneconomic level, and we decided that we would simply establish that Federal Government price, and stick to it, which is what we did', ibid., Part 24, p. 13694. It may be doubted that 'uneconomic levels' implied losses. On the basis of the cost data obtained by the Subcommittee, and assuming that Pfizer paid the same royalty on Lederle's process that Bristol did, that Lederle's and Pfizer's production costs were somewhat lower than Bristol's, and that Squibb's costs were the same as Upjohn's, Lederle's production costs and royalties might be $3.60; Pfizer's, $4.00; Bristol's, $5.03, Squibb's and Upjohn's, $9.30. This might aid in explaining the lower initial bids of Lederle and Pfizer.

[74] Testimony of Commander Arnold Wales, U.S. Navy Supply Corps, ibid., Part 24, p. 13815.

incidence of excess capacity or excess inventories, but does not reflect either relative production or relative sales. In 1958, Bristol produced 36 per cent, Lederle 33, and Pfizer 31. Sales are not available, but Bristol sold about one-third of its output to Upjohn and another third to Squibb,[75] so that relative sales might be: Lederle, 33 per cent; Pfizer, 31; Squibb, 13; Bristol, 12; and Upjohn, 11, unless Pfizer purchased significant quantities from Lederle, as it is entitled to do by agreement. It seems likely that, although prices fluctuated, real price competition never entered this market until foreign firms, not encumbered by restrictive patent agreements, were permitted to bid.[76]

IV. CONCLUSIONS AND POLICY RECOMMENDATIONS

Evidence has been presented to show that effective price competition among ethical drugs is seriously limited by the patent privilege. Holders of patents (or sometimes merely of patent applications) may legally exercise restrictions on output and maintain prices at levels that are extremely high relative to production costs. The resulting gross profit margins are employed in large measure to finance enormous advertising and sales promotion campaigns which contribute materially to the already grave imperfections of market information. By this means, small sellers of generic name drugs are deprived of the physician's attention, and cannot obtain any significant share of the prescription market, even though they may be selling at prices which are a small fraction of their larger rivals'. This applies not only to products protected by patents or patent applications, but to virtually all advertised drugs. Monopoly profits from the sale of patented drugs thus finance advertising campaigns which extend monopoly power into other drug markets which may not be protected by patents.

Drug firm monopoly and oligopoly could perhaps be rather readily supplanted by workable competition if two simple but radical reforms were effected in the institutional structure of the drug market: the abolition of the patent privilege as it applies to drug products, and the expansion of the powers of the Food and Drug

[75] Data supplied by Bristol, ibid., Part 24, p. 13907.

[76] Admiral Knickerbocker offered some interesting testimony which may tend to illuminate the chiaroscuro pattern of competition in the bid market from a somewhat different perspective. The decision of the Military Medical Supply Agency to accept foreign bids was not reached until after a meeting with Pfizer representatives, during which, Admiral Knickerbocker testified, '. . . Mr. Cooney [Pfizer's sales representative], in rather an unguarded moment, made the statement that the price of tetracycline would stay where it was until Pfizer did something about it', ibid., Part 24, p. 13819. Had the network of patent control been absolutely worldwide, this alleged 'unguarded moment' might not have cost Pfizer any sales.

Administration to insure fully adequate drug plant inspection and control of drug advertising.

It is something of an anomaly that the United States patent laws permit drug product patents; about two-thirds of the countries in the world prohibit them on grounds of public policy, and all but three of those which do permit such patents have provisions for compulsory licensing. Drug product patents should be abolished; drug process patents should be subject to compulsory licensing at reasonable royalties. This may or may not reduce the volume of research done by private drug firms. In turn, a possible diversion of drug research effort from private to public channels may or may not improve the efficiency of resource allocation in this sector. In many foreign countries, drug research has been immensely productive in the absence of drug patents. In the United States, patent protection has allowed prices to be maintained at very high levels and has detrimentally affected the nature of research. With prices enormously above costs (defined as raw material cost plus production cost plus a competitive rate of profit on the investment in productive facilities), the pressure to obtain more volume is inevitably diverted into lavish selling expenses, and into the 'research' of molecular manipulation.[77] Basic research by drug firms may be of questionable productivity, but the high salaries paid absorb all too large a fraction of those very scarce human resources qualified to engage in basic biochemical and pharmacological research. It is common knowledge in the industry that each major firm's research programs duplicate those of their rivals; witness the near-simultaneous discovery of many antibiotics and corticosteroid hormones by two or more firms. The total productivity of drug research efforts would arguably be increased by a partial shift from private to public and university channels.

The Food and Drug Administration should be given sufficient additional funds to allow adequate inspection of all drug producers, large and small. It should also be given the authority to regulate all drug advertising and labelling, with a view toward eliminating brand names in favor of identification by generic name plus the name of the seller.

Such reforms, properly implemented, might for the first time

[77] Abundant evidence was presented during the Senate Hearings to show that much drug product 'research' is misnamed, and could more appropriately be referred to as product development and product promotion. Dr. A. D. Console, former medical director of Squibb, testified to this effect, adding: 'I think the majority of it [drug firm research] is in that category . . . with many of these products, it is clear while they are on the drawing board that they promise no utility. They promise sales', ibid., Part 18, p. 10380. See also the testimony of Dr. H. J. Weinstein, former medical director of Pfizer, ibid., Part 18, pp. 10243–4

make the ethical drug market workably competitive. The elimination of patents would (in the absence of economies of scale in production) lead to the entry of new firms producing under conditions of approximately constant costs. The inelasticity of demand would thus be offset by highly elastic supply. The existing economies of scale in selling would, in time, be eliminated by advertising reforms and the disappearance of those monopoly profits which have motivated and financed extravagant selling efforts. The advertising reforms would also remedy the imperfection of market information, allowing physicians to become aware of the precise nature of the price and product alternatives, and disparagement efforts would wane as the funds necessary to finance them dwindled due to increasing competitive pressure on prices, and as the manifest adequacy of FDA inspection rendered the substance of such contentions obviously specious.

RICE UNIVERSITY
HOUSTON, TEXAS

THE FORTUNES OF ECONOMIC REFORM LEGISLATION: THE CASE OF THE DRUG AMENDMENTS ACT OF 1962

(By Henry Steele[1])

I. INTRODUCTION

On April 12, 1961, the late Senator Estes Kefauver introduced a bill (S. 1552, 87th Congress, first session) to amend the antitrust and food and drug laws with regard to the prescription drugs industry. This bill, the outcome of extensive hearings held by the Subcommittee on Antitrust and Monopoly of the Senate Judiciary Committee in 1959–1962 regarding administered prices in the drug industry, embodied provisions to prevent monopolistic abuses in drug marketing. The Report issued by the Subcommittee established the existence of considerable monopoly and oligopoly power on the part of major drug firms, and indicated the extent to which such market power was obtained by virtue of patent privileges, excessive advertising, and other marketing practices which tended to augment the imperfection of market information in the interests of the larger firms.[2]

II. ANALYSIS OF DRUG MARKETING PRACTICES

An economist would briefly summarize the drug marketing situation as follows: demand for many prescription drugs is almost perfectly inelastic. Supply can be restricted by drug firms which secure product patents and refuse to license them. The resulting enormous profit margins on such products (factory costs may be less than 10 percent of wholesale price[3]) stimulates entry, which take the form of imitative research activity on the part of other drug firms, aimed at devising a substitute for the patented drug which they can patent and advertise as superior to the original. Rivalry is thus diverted from price competition to product differentiation, as rival patented drugs attempt to take over the prescription market for a given group of disorders by intensive advertisement to doctors (the only agents qualified to prescribe) through direct mail, itinerant salesman (detailmen), and the sponsoring of medical conventions. High unit profit margins can finance large (and often imitative, duplicative, and wasteful) research programs, overpowering advertising campaigns, and still yield very high profits on investment. Since costs are low, and economies of large scale production apparently unimportant, small firms might be able to compete with large firms on a cost basis, but they suffer from two disabilities. Being unable to finance sales campaigns, their products do not come to the attention of the prescribing physician. Furthermore, the detailmen employed by the major firms make a practice of disparaging low-priced drugs, inducing the physician to equate low price with low quality. At best, the physician may not be very price-conscious, since he does not pay for the drugs he prescribes.

The situation is aggravated by weaknesses in patent laws and practices, and in the regulations governing drug advertising. Patent protection is absolute as regards a given compound, but weak as regards slight variations from that compound. The game of "molecular manipulation" is a popular one with drug firm research personnel: the goal is to devise a derivate compound which has the same therapeutic effect as some primary patented compound, but which by virtue of its marginally different chemical structure can be patented as a different drug and advertised as a "new and superior" healing agent.[4] Patents can also be obtained for naturally occurring substances if never previously isolated, for processes occurring in nature, such as fermentation, and even for combinations of existing patented drugs. But since most major drug firms have research programs which substantially duplicate each other, near-simultaneous discovery of the same drug is a not infrequent occurrence. When several patent applications covering the same compound are received by the patent office, an interference is declared, and the parties attempt to assist in resolving the question of priority to the satisfaction of the patent office. There is, however,

---

[1] The author is assistant professor of economics at Rice University.

[2] *Study of Administered Prices in The Drug Industry, Report of the Subcommittee on Antitrust and Monopoly of the Senate Judiciary Committee, Pursuant to Senate Resolution 52, Eighty-Seventh Congress, First Session,* Washington, D.C., Government Printing Office, 1961.

[3] *Ibid.*, pp. 15–16.

[4] *Ibid.*, p. 16.

no barrier to private settlement of interferences by agreements among the parties themselves, involving concession of the patent to one firm in return for granting licensing rights and perhaps other privileges to the other firms. The patent office rather welcomes such private settlements in view of the pressure of work on its staff, but the antitrust division has been critical of the nature of many private interference settlements.[5]

Advertising abuses relate chiefly to drug nomenclature and to the quantity and quality of sales efforts. Four hundred or more new drugs are marketed each year in the United States. Each new drug should be given a generic name to identify it; individual firms marketing the drug may confer upon it their own brand names. If a physician prescribes a drug by generic name, the pharmacist may fill the prescription with any brand of the drug; if the prescription is by brand name, only that firm's brand may be dispensed. Evidence indicates that generic names are designed to minimize use (excessive length, complexity, and clumsiness) while brand names are brief and memorable. Regulations require that generic names be displayed in all advertisements, by they are often obscured by the use of microscopic type face, by being concealed in obscure places in the advertisements, and are sometime simply omitted.[6] Some drugs have no generic name; others have more than one. This deliverately cultivated confusion is evidently intended to suppress use of generic names in favor of brand names.[7] By this means, trademarks are made to supplement patents as monopolistic devices. The same patented drug may be sold by ten different licensees under ten different highly advertised brand names, and each licensee through sales efforts may succeed in differentiating in the mind of the physician this physically homogeneous drug. In those rare cases where there is no patent protection, generically-named products sold by small firms may be priced at ten per cent or less of the price of the major drug firms, but price competition is rendered impossible unless prescriptions are written generically. However, the detailmen have been very successful in their disparagement campaigns; surveys show that almost 90 per cent of drug prescriptions are written by use of brand names.[8] The quantity of drug advertisements is overwhelming in itself, including not only propaganda but also free drug samples, trinkets (toy urinals, chinese dolls, head cushions, etc.) and gifts such as golf balls engraved with the recipient's name. But the quality of advertising claims is the greatest obstacle to the enlightened practice of medicine. Dozens of examples of misrepresentation were unearthed at the Senate hearings; one example must suffice: a drug firm mailed to physicians advertising copy showing X-ray photos clearly designed to imply a dramatic recovery in a patient's condition before and after the use of the advertised drug. Upon inquiry to the firm's medical director, it developed that the two X-rays were of entirely different persons with qualitatively different disease conditions, and that neither had ever used the drug being advertised.[9]

III. INDICATED REFORMS

To an economist, the abuses in drugs seem amenable to relatively simple reforms. Drug product patents should be abolished; drug process patents should be subject to compulsory licensing at reasonable royalties. Generic names should be simplified, the use of brand names should be eliminated, and firms should be required to advertise and sell their products under labels giving the generic name of the drug, followed by the name of the firm, *e.g.*, chloramphenicol-Parke, Davis instead of "Chloromycetin." Food and Drug Administration drug plant inspection authority and funds should be increased in order to guarantee the safety of all drugs on the market, rendering specious all disparagement campaigns. Control of the quality of drug advertising should be made truly effective. If rigorously enforced, these reforms should suffice. Absence of patent protection and mandatory use of generic names would allow price competition between large and small firms; excessive profits would disappear, and with them would disappear the ability to carry out wasteful and duplicative "research" programs,

[5] *Ibid.*, p. 48.
[6] *Study of Administered Prices, op. cit.*, p. 234.
[7] Firms may, however, on occasion find generic names of use as when the necessity arises for sending out circulars to warm physicians about side effects. *Hearings on S. 1552*, Part 5, p. 2938.
[8] *Ibid.*, Part 18, pp. 10481–10482.
[9] Ibid., Part 7, pp. 3301–3310.

and to advertise so intensively as to prevent the products of smaller firms from coming to the attention of the physician. Much of drug research would thereby be shifted from private to public and university channels, with a probable increase in the efficiency of such efforts, since the patent-induced incentives to engage in duplicative research, molecular manipulation, and promote unpromising drugs in the interests of their sales potential rather than their efficacy, would be removed.

IV. S. 1552: FROM BILL TO ACT

A. Original provisions

Senator Kefauver, however, did not contemplate so sweeping a reform of existing practices, perhaps in view of the political difficulties. His original bill did, nevertheless, incorporate a number of valuable reforms, particularly of patents and drug nomenclature. The salient provisions of the bill may be listed as follows:

(1) Drug patents shall take effect as of the date of the new drug application (or in the case of drugs for which no new drug application is necessary, the date for filing the patent application) and shall grant unconditional monopoly power for a period of only three years, after which such patents are subject to compulsory licensings at a royalty not exceeding 8 per cent of the sales price.

(2) No patents are to be issued for compounds merely embodying molecular manipulations of existing drugs, or combinations of existing drugs, unless the Secretary of the Department of Health, Education and Welfare determines that such a drug has a significantly greater therapeutic effect than the unmodified drug, or of the combined drugs when taken separately.

(3) The Sherman Act should prohibit private drug patent interference settlements or any other private arrangements whereby parties concede patent priority claims in return for royalty-splitting agreements, differentially favorable royalty rates for participants relative to third parties, or limitation of licensing to parties to the agreement.

(4) The Food and Drug Administration should be given authority to establish generic names of drugs, with a view toward simplification and increased use of such names.

(5) The Food and Drug Administration shall pass upon the efficacy as well as safety of drugs.

(6) The Secretary of the Department of Health, Education and Welfare should publish and distribute to physicians and hospitals copies of drug firm package inserts which describe the action of drugs and give data on dosage, contraindications, and side effects. (At present, such inserts go by only to pharmacists, not physicians.)

(7) All drug firms must be licensed by the Secretary of the Department of Health, Education and Welfare, and must submit to inspections of plant and equipment. (Plant inspection regulations are also to be made more adequate.)

(8) All drug advertising shall include (a) the generic name in type face as large as the brand name, (b) a statement of the conditions for which the drug is an effective treatment, and (c) a statement of all side effects and other warnings.

The reforms contemplated are rather modest. The provision for making patents effective as of the date of filing a new drug application is intended to eliminate any incentive for delay on the part of the applicant in patent proceedings, particularly in negotiating interferences in the interest of prolonging effective patent protection, not only for the successful applicant, but also for the entire group of interim licensees. The three-year term of absolute patent monopoly is intended to allow the successful firm to recoup its research costs during an initial monopoly period of high prices and no rivals; the compulsory licensing provision during the next 14 years is tempered by allowing the patent holder to charge an 8 per cent royalty, a very high royalty by drug industry standards, the usual rates being between two and six per cent. The prohibition of drug combination or molecular manipulation patents in the absence of evidence of superior efficacy is a laudable (although administratively difficult) attempt to halt wasteful duplicative research efforts.

The proposed Sherman Act amendment is rather strict in that it entirely prohibits private drug patent interference settlements, by making it illegal to withdraw any pending drug patent application. This is desirable in that private settle-

ments in effect transfer the determination of priority of invention from the Patent Office to the rival parties themselves. Such private agreements often contain stipulations which restrain trade. One party is awarded the patent, and as a condition of surrendering their claims, the other parties are guaranteed licenses, and may even begin to pay royalties before the patent is issued. The agreements usually contain restrictive provisions under which the parties agree not to sell to outsiders, not to sell the drug in bulk powder form (which might get into the hands of non-licensed dealers who could tablet and bottle it and sell it at low prices), and to accept other limitations on their marketing practices. In addition, all of the parties to such agreements will almost invariably sell at identical prices.[10]

Most of the other proposed reforms would alleviate abuses in drug marketing. Although the use of brand names is not outlawed, reforms in generic names would be a step in the right direction. The requirement that the FDA pass on the efficacy as well as on the safety of drugs is a very reasonable one, but one which was bitterly fought by some drug firms during the original hearings on administered prices. In the absence of such a requirement, drugs which are not clearly harmful, and for which various physicians have written "testimonials" supporting a new drug application, may be allowed on the market, at which point the market success of the drug depends (at least for a while) more on the skill of the sales department than on the intrinsic therapeutic merits of the compound. Several physicians testified at the hearings that not a few drugs currently on the market are absolutely useless.[11] The sixth provision is aimed at one of the many paradoxes in drug marketing. Drug makers are required by FDA regulations to print up a leaflet describing the uses, dosages, and side effects of each prescription drug, but the leaflet need only be included in the drug package, which goes to the pharmacist, who, having no use for it, throws it away. The physician, to whom such information is of the most vital importance, has no ready access to this information, but must routinely rely on advertisements and detailmen. Distribution of all such material to physicians is obviously an imperative necessity. Finally, the provisions for licensing all drug makers and for requiring more adequate inspection of drug plants is intended to insure the quality of all drugs, generic and brand name, and thus hopefully to reduce the effectiveness of generic drug disparagement and increase the physician's willingness to prescribe by generic name.

B. Senate hearings on S. 1552

The hearings on S. 1552 were held between July 1961 and February 1962. The battle lines were rather tightly drawn up on most issues.[12] In the medical profession, the American Medical Association (AMA) betrayed dedicated opposition to every section of the bill upon which they took an official position; on the other hand, of the eleven medical educators who appeared either individually, or as representatives of such groups as the American Public Health Association, ten favored the proposed reforms and an eleventh expressed opposition to the patent provisions, an area perhaps not strictly within his professional competence. Other physicians appeared in their capacities as administrators of hospitals or health insurance plans, and indicated general approval. Ten patent attorneys and other patent spokesmen testified; nine of them, including representatives of the American Bar Association, The National Association of Manufacturers, the American Patent Law Association, *etc.*, found little or no virtue in any part of the bill. The tenth, an attorney from a small town in a mid-western state, and representing only himself, regisitered substantial agreement with the aims of the bill. The Pharmaceutical Manufacturers Association opposed most of the bill's provisions, but showed limited agreement in some areas; other drug industry witnesses were not nearly so favorable.

Testimony was given by two economics professors. One appeared at the request of the Subcommittee as an expert on the economics of patents, and gave testimony to the effect that the patent provisions of S. 1552 would probably not cause

---

[10] *Report on S. 1552*, op. cit., p. 45.

[11] See, for example, *Hearings, op. cit.*, Part 1, p. 285.

[12] Battle lines were also drawn up within the Subcommittee itself, the majority (often represented by Kefauver alone) in favor of reform, and the minority (represented by Dirksen and Hruska) apparently concerned with obstruction of the proceedings and harrassment of witnesses favorable to the bill. Hruska in particular was adept at the latter, not scrupling to insinuate Communist leanings in those with whom he disagreed. See *Hearings, op. cit.*, Part 3, pp. 1410–1411.

the ruin of the drug industry if adopted; indeed, he saw many faults in the present patent system, and suspected it would be possible to increase the efficiency of drug research by transferring some of it from private to public auspices. The other was retained by counsel for the Pharmaceutical Manufacturers Association, and he could find no evidence of any need for Kefauver's bill.

Testimony was also obtained from a number of labor and consumer groups. Representatives from the AFL–CIO, the UAW, and the International Union of Electrical, Radio and Machine Workers supported the bill, as did spokesmen for Consumers Union, the National Consumers League, and the Cooperative League of the United States. Two representatives of retail pharmacy appeared. One spoke for the American Pharmaceutical Association, a group of pharmacists. He articulated a viewpoint consonant with that of the drug makers, and expressed great concern that the government not take the unwise step of distributing drug information to physicians. Another pharmacist, a former teacher and state pharmacy and drug law enforcement officer, found considerable merit in many of the reforms. The last parties to testify were advertising agency representatives who predictably saw no apparent need for any advertising reforms.

It is likely that the most influential testimony was that given by the American Medical Association, the various patent law groups, and the Pharmaceutical Manufacturers Association. It is instructive briefly to review the character and merits of the testimony given by these groups.

1. Testimony of the American Medical Association

The AMA made no recommendations in regard to the antitrust law amendments, but took the view that none of the other proposed reforms were defensible, including the requirement that the FDA pass on the efficacy of drugs. Their spokesman, Dr. Hugh Hussey, recognized the need for certain improvements in regard to drug nomenclature and physician information. The AMA preferred , however, to carry out the reforms itself, in cooperation with the drug industry, but with no participation by any public bodies.[13] One may be forgiven for entertaining the view that the AMA position is simply a defense of the revenues it obtains from drug firms for advertisements in its journals. The roots of AMA opposition are more ramified and complex, but it is easy to trace the influence of advertising revenues, as seen against the background of prior AMA actions.

First, the AMA reform program was adopted only five weeks prior to the scheduled appearance of its representatives before the Subcommittee, and the implementation of the program was to be gradual, extending over two or more years. Five years previously, however, a similar reform program in response to similar legislative demands, was outlined by the AMA, proposing cooperation between itself and the drug industry to control misrepresentation in advertising, but it entirely failed of implementation.[14] Second, the AMA has become increasingly dependent upon drug advertising for its own financing. In 1949, medical journal advertising revenues comprised about 31 per cent of all AMA revenues: in 1955, about 44 per cent; and in 1960, a little over 50 per cent. Total advertising revenues, however, been augmented by royalties received from the leasing of the rights to use mailing lists of physicians.[15] The sums received have increased from small amounts to about 5.6 per cent of total revenues by 1960. Hence in 1960 about 56 per cent of AMA revenues came from drug firm advertising efforts. Third, the AMA in recent years has become increasingly permissive in its attitudes toward advertising standards, and it is likely that the increase in its advertising revenues is in good part attributable to this. The period of increasing leniency coincided with that of increasing advertising revenues. The measures by which this more lenient policy was evolved, or from which it can be inferred, may best be described and interpreted in chronological order.

---

[13] *Ibid.*, Part 1, pp. 47–49.

[14] *Ibid.*, Part 1, p. 341. Testimony of Dr. Allan M. Butler, Professor Emeritus at Harvard University.

[15] Obviously a royalty on direct mail advertisements ($2 per thousand mailings) creates a direct financial interest in maximizing the volume of such traffic, and is hence undesirable in even greater degree than policies to increase advertising in the AMA journals. There is evidence that advertising standards in AMA journals are still higher than those in most (but not all) other medical journals, but there is no AMA control over the quality of direct mailings. The royalty income feature of AMA finances tends to justify in part the accusation of James Carey, president of the International Union of Electrical, Radio, and Machine Workers, that "The AMA, in our opinion, is just a business." (*Ibid.* Part 5, p. 2731.) It is ironic that, while most students of the AMA compare it to a trade union, a leading trade union spokesman sees it as a business.

In 1950, a rule (adopted in 1905) that only the inventor of a new drug could use his brand name of the drug in advertising in the AMA journals, was dropped. In 1952 the publication of a handbook, *Useful Drugs* (issued periodically since 1917) was discontinued. (The 1961 AMA proposals contemplated the publication of a similar book which, however, would include all drugs and would thus lack the discrimination between useful and less than useful drugs characteristic of the earlier publication.) In 1955, major advertising policy changes were made in response to a survey by Ben Gaffin and Associates addressed to the problem of increasing AMA advertising revenues. Before 1955, the AMA Council on Drugs, an expert technical body, had effective control over advertising. Only Council-approved drugs could be advertised. Approval required the obtaining of a "Seal of Acceptance" which was granted only after the drug firm had submitted to the Council all the data it requested, including both favorable and unfavorable reports. (By way of contrast, at that time there was no requirement that all test data, favorable and unfavorable, be sent to the FDA with new drug applications.) The Council on occasion inspected drug factories. It exercised some control over generic names, since only generic names approved by the Council could be used in AMA journal advertising. The Council examined and passed on all advertising.

In 1952, Ben Gaffin and Associates began a survey to determine how AMA journals could get more advertising. It noted that AMA advertising had increased only 3 per cent since 1948, while other medical journals had increases of 40 per cent. The survey indicated that while physicians and small drug firms thought well of AMA advertising control policies, large firms were critical, particularly of the "Seal of Acceptance" program. The Gaffin study recommended that the Seal of Acceptance program be dropped, and the advertising controls be liberalized.[16]

In 1955, several major changes were made in advertising policy. Advertising control was taken out of the hands of the Council on Drugs. The Seal program (in effect in that form since 1929) was abolished. The Council lost its influence over generic names by this means, and also its power to elicit unfavorable as well as favorable evidence on drugs. The Council lost all control over advertising.[17] While rules requiring generic name advertising were being abolished, editorial changes were made which drew increased attention to brand names.[18] It is instructive to note that the 1961 reform proposals of the AMA did not include reinstatement of the control over drug advertising on the part of the Council on Drugs. Indeed, the preparing its presentation for the hearing, it did not even seek the advice of the Council, its own expert advisory body on the subject matter to which the hearings were addressed. This fact was testified to with dismay by more than one member of the Council.[19] Had the Council been influential in drafting the AMA's testimony, it is hardly conceivable that the AMA would have gone on record as opposing the requirement that a drug need be efficacious in order to merit a new drug permit.

During 1955 a microbiological laboratory, established in 1949 to establish purity standards for antibiotics and other drugs, was abandoned. And in 1959 a chemical laboratory, founded in 1906 to formulate and develope standards for drugs, was abandoned because of overwork and lack of finances.[20]

---

[16] *Ibid.*, Part 1, pp. 102–103; Part 2, pp. 490ff. Part II of the survey includes the following: ". . . while the possibility of increasing advertising revenue by several million dollars per year is a good motive for putting into effect the information gained from these two studies, there is an even more important reason for so doing. . . . the AMA has an opportunity to assume leadership in improving some $130 million worth of medical advertising per year." It thus becomes apparent that it was the duty of the AMA to enrich itself.

[17] Mr. Stetler of the AMA denied that these changes had occurred in response to Gaffin's survey. Gaffin himself was convinced otherwise. In another drug survey, he links his recommendations and the AMA's actions: "The survey of pharmaceutical advertisers played a part in bringing about a number of policy changes . . . [including] the eventual dropping of the 58-year old council seal of acceptance program." *Ibid.*, Part 1, p. 125. If the AMA disagreed with Gaffin's diagnosis, they took no disciplinary action. Gaffin continued to conduct their surveys throughout the 1950's.

[18] Dr. Hussey volunteered that "Trade names, formerly listed at the end of monographs, also have been placed immediately after the nonproprietary titles of monographs and on the front page of the *Journal* to more readily catch the eye of *Journal* readers." *Ibid.*, Part 1, p. 106.

[19] *Ibid.*, Part 1, pp. 216, 375.

[20] The Council of Drugs also suffered from insufficient finances, its appropriations increasing from $135,000 in 1950 to $156,000 in 1954 (the last year of active Council advertising control) and declining to $75,000 in 1960. Curiously, the public relations department budget increased prodigiously, from $102,000 in 1950 to $494,000 in 1958. *Ibid.*, Part 1, pp. 126–127.

It seems safe to conclude that on the basis of the recent record, the role of the AMA in contributing to high standards of drug advertising has steadily diminished. The adoption of the joint AMA-drug industry program outlined by Dr. Hussey would provide at best a questionable safeguard for the drug consumer.

2. Witnesses Testifying on Patent Provisions

The great majority of the testimony on the proposed patent reforms revealed the presence of a tropismic conservative reaction against any modification of the patent system, and the absence of any great evidence that the witnesses had studied the concrete operation of the patent system in the framework of the drug industry.[20a] The presentation given by Joseph Jackson, the chairman of the patent, trademark, and copyright law section of the American Bar Association, may serve as an illustration. Speaking for the American Bar Association (ABA), Jackson reported that the ABA Board of Governors had adopted five resolutions in regard to S. 1552: (1) disapproval of the antitrust provisions limiting private patent interference settlements; (2) disapproval of the requirement of "non-obviousness" regarding the patentability of a drug product; (3) disapproval of the requirement of proof of significantly greater therapeutic effect in order to qualify molecular modifications for drug patents; (4) disapproval of the distinctive treatment of drug patents with regard to the date of effective patent protection; (5) disapproval of compulsory licensing for drug patents.[21]

When Mr. Jackson was examined on his testimony, it developed that these resolutions, purporting to speak for the ABA, a body with a membership of some 102,000, had been drafted in Saint Louis, about ten days before the Hearings, by a group of 150 to 200 patent lawyers, with orders to "act with great acceleration" because of the "emergency." [22] These resolutions were then submitted to the Board of Governors of the ABA (no member of which is a patent lawyer) and were promulgated by them more or less over the head of the House of Delegates, the representative deliberative body of the ABA. (This is an extraordinary procedure which is nevertheless permitted by the constitution of the ABA.) Furthermore, members of the ABA were apparently not given notice of this action by their Board of Governors, as evidenced by the surprise of several members of the Subcommittee staff belonging to ABA. It further developed that the ABA group required less than an hour and a half to reach conclusions on matters which the Subcommittee and its staff had been studying for over two years. Jackson admitted that there had been no discussion of the economics of the drug industry and its relation to patents, no study of costs, no systematic consideration of profits, no attention to concentration or to the interdependence of major firms, no concern with entry conditions, no inquiry into price policies, and not even any consideration of the nature of incentives and the quality of research effort in drugs. Jackson explained that none of these matters had been considered, since "If we had discussed all these subjects you are presenting, we would still be in Saint Louis. We would not be here with any resolutions at all." [23] But that is precisely the point. It may be inferred that the only aim of the ABA action was to go on record as condemning S. 1552 in time for the Hearings, regardless of the factual merits of the bill.

The patent provisions stand or fall depending upon their application to the specific circumstances of the drug industry; the entire patent system is not at stake. But patent attorneys in testifying almost universally took the position that any amendment to the patent laws in respect of a particular industry would necessarily imperil the patent laws with regard to all other industries. Jackson no doubt faithfully reported this attitude on the part of the ABA group when, in response to an observation by the Subcommittee counsel that S. 1552 was limited in its application to the drug industry alone, "We were afraid we would be faced with a special antitrust law and a special patent law for butter and eggs, and another one for milk and beer, and so on in different areas." [24] One gets

[20a] There is no evidence of any articulate "grass roots" support for the opponents of the patent provisions of S. 1552. Symptomatic of the testimony in behalf of the patent *status quo* is a communication from an individual styled "Clair V. Johnson, Newfane, Vermont, Patent Lawyer," who roundly condemns the bill. It appears that Mr. Johnson is also a director of U.S. Vitamin and Pharmaceutical Corporation. (Part 3, p. 1601).

[21] *Ibid.*, Part 3, p. 1473.

[22] *Ibid.*, Part 3, p. 1472.

[23] *Ibid.*, Part 3, p. 1481.

[24] *Ibid.*, Part 3, p. 1480.

the impression from the main drift of the patent testimony that this rather naive contention is not advanced merely as a straw man, but that this consideration alone was enough to close the minds of patent attorneys regarding S. 1552. A further observation by Jackson leads one to wonder why his group should have debated even as long as an hour and a half on their resolutions: "There was only one man of the whole group who had anything good to say about the law . . . he said it might not be politically expedient to show the full extent of our disapproval." [25]

Had the ABA group referred to the Subcommittee's *Report* on the drug industry, they would have discovered that it is not at all unusual for nations to make exceptions of their drug industries in regard to patent privileges. Of some 77 nations with patent laws, 49 absolutely prohibit drug product patents on grounds of public policy, and 25 others have provisions for compulsory licensing. Only Panama, Belgium, and the United States allow unrestricted drug product patent privileges.[26] Indeed, to judge by a comparison of drug discoveries by drug firms in countries with and without product patents, it is by no means clear that the patent incentive is necessary to elicit productive drug research. The patent privilege restrictions embodied in S. 1552, although more liberal, are closely related to those of Germany, long one of the world leaders in drug research. German patent law denies drug product patents, but drug processes may be patented, and such patents cover the products made by those processes. If, however, alternative processes are devised to produce the same drug, the drug in question then fails to retain its effective protection. Germany permits unrestricted drug process patents for a period of three years. After that, compulsory licensing is required, with royalties of between 5 and 10 percent, as determined by the decision of a special tribunal.[27] Professor Machlup of Princeton University testified that not only has the existence of such a patent law in Germany failed to halt productive drug research, but that its expediency and equity is not even questioned.[28]

One ominous tendency which came to light at the patent hearings concerns certain evidence that pressure groups are attempting to weaken the protection which drug buyers currently enjoy under the patent laws of many nations of the world. Professor Machlup testified that in recent years several countries, in consequence of pressure by industrial groups, have made their patent laws more favorable to such industries.[29] It is ironic that at the time when efforts are being made to bring United States drug patent policy into line with the more enlightened practices of other industrial countries, some of these very countries are experiencing a retrograde tendency. In 1949, England amended its patent law to allow drug products to be patented, but required compulsory licensing. France adopted a largely similar law in 1960.[30]

Drug and other chemical interests appear to have been active in connection with a certain diplomatic conference held in Lisbon in 1958 for the purpose of revising the International Convention for the Protection of Industrial Property. Prominent among the United States delegation were Roland Libonati, a Representative from New Jersey (where many drug firms have plants) and P. J. Federico, examiner in chief of the United States Patent Office, who testified at the hearings and was at pains to take issue with the Subcommittee's contentions on the relative strength of patent protection in countries with and without product patents. The Lisbon convention agenda contained an item proposing to require all countries adhering to the Convention to grant patents for chemical products, including pharmaceuticals. The United States delegation sponsored this resolution, and it fell to Mr. Federico's lot to expedite the proceedings. The resolution did not pass—12 nations voted against it—but observers were somewhat surprised to find that those voting in favor included the delegations from eight countries which prohibited drug product patents. After the resolution failed, Germany introduced a resolution to recommend that member countries study the question with a view toward revising their patent laws. This resolution passed. Such efforts at "study" may be bearing fruit: France, one of the countries opposing the original resolution in 1958, adopted its own drug product patent law in 1960, and the Scandinavian countries and Finland are trying to work out a common patent law which will extend to drug products.

---

[25] *Ibid.*, Part 3, p. 1480.
[26] *Study of Administered Prices, op. cit.*, p. 106.
[27] *Hearings on S. 1552, op. cit.*, p. 1385.
[28] *Ibid.*, p. 1385.
[29] *Ibid.*, Part 3, p. 1371.
[30] *Ibid.*, Part 3, pp. 1201, 1216.

3. Testimony in Behalf of the Pharmaceutical Manufacturers Association

The influence of the drug makers lobby in behalf of aborting reform efforts was not limited to their presentation of testimony at the hearings, although a survey of their manifold efforts a influencing—and even drafting—the legislation on to control their own activities is beyond the scope of this paper.[30a] Suffice it to say that the legislators must have found the PMA testimony the most convincing case with which they were presented, since the PMA recommendations agreed to a great extent with the Act as it was finally passed. Mr. E. N. Beesley, chairman of the PMA, presented his views as follows: (1) agreement with the the provision that FDA should pass on the efficacy of drugs, but judged on the basis of "substantial" and not "preponderant" evidence; (2) agreement that distribution to physicians of full information on drugs is important, but that the provisions of S. 1552 in this respect are unnecessary, in view of the AMA's promotion of a cooperative program with PMA which will handle these details; (3) qualified agreement that some more systematic approach to the designation of generic names is desirable, but a preference for the announced AMA program of voluntary cooperation with PMA on generic names, with the concession that if such voluntary means do not achieve results, the FDA may designate a generic name; (4) disagreement with the compulsory licensing of drug manufacturers, but a readiness to compromise on simple registration, together with an expanded program of factory inspection by the FDA; (5) complete rejection of all phases of the patent program: condemnation of compulsory patent licensing, disapproval of the restrictions of the patenting on molecular modifications and drug combinations, and disagreement with the contemplated restrictions on private patent interference settlements, but with the concession that (as suggested by Patent Commissioner Ladd and Assistant Attorney General Loevinger) firms be required to file private patent interference settlements and other relevant data at the Patent Office for public inspection.[31]

The PMA testimony revealed a supporting willingness to compromise in regard to some parts of the bill, but the economic impact of these concessions would be but slight. It is also to be noted that while much is made of the prospects for cooperative programs with the AMA, nothing is included which would bring advertising claims under any direct control. The prospects for constructive reform in these respects are made doubtful not only by the past record of the AMA but also by certain recent activities of organized drug makers in relation to attempts by a group of physicians, acting independently, to increase the standard of drug advertising.[32]

The PMA showed a clear intent to resist economic reforms which would reduce drug profits. Mr. Beesley testified directly to this effect: 'Deletion of patent and trademark protection . . . would increasingly dilute the intensive competition

---

[30a] Kefauver, angered at the opposition of the subcommittee minority, and apparently suspecting it of connivance with the drug lobby, took the politically risky step of confiding in a journalist, Richard Harris, who was then doing a study of the drug industry. Kefauver made a detailed first hand report of the day by day political maneuvering. The resulting chronicle, summarized in the New Yorker, March 28, 1964, pp. 46 ff., makes fascinating reading. It appears that PMA's initial reaction to Kefauver's bill was divided. One faction reportedly preferred to buy outright the necessary votes to defeat the bill; another thought it wiser to influence legislation in the making, conceding some points in order to moderate reformist pressures and prevent the eventual passage of truly substantial reforms. It appears that, under White House auspices, representatives of PMA were asked personally to aid in drawing up an alternative drug bill to replace that of Kefauver. At the subsequently notorious "secret meeting" at which the substitute bill was drawn up, the only persons present were PMA lobbyists, legislative draftsmen from the Department of Health, Education, and Welfare, and counsel for the Subcommittee minority. Not a single senator was present. Nor had Kefauver, or any other member of the subcommittee majority been informed of the meeting. Fortunately, this attempt to railroad through an impotent bill was nullified when Kefauver succeeded in drawing public attention to it and the Administration saw fit to repudiate the substitute bill.

[31] *Ibid.*, Part 4, pp. 1997–2007.

[32] The AMA seal of acceptance had been valued by some of the more responsible drug firms, large and small. With the demise of this program, some firms then supported a group of physicians who undertook to establish drug advertising standards and to review advertising copy, and who later became known as the Physicians Council. This group awarded approved advertisers with an insignia which could be displayed in advertisements. As the program became more widely accepted, the industry began to view it with suspicion, and in February 1959 the medical section of the American Drug Manufacturers Association (now part of PMA) adopted a resolution calling upon its members to refuse cooperation with the Physicians Council. The resolution claims that such outside review ". . . would disrupt the normal, prompt, and effective processes of ethical drug promotion," a statement with which the Physicians Council would probably not find fault. The financial support of the Council disappeared, and they abandoned their insignia program. Testimony of Dr. Julius Richmond of New York State University, *Ibid.*, pp. 361–368.

for superiority in discovery and manufacturing; it would leave price competition as the only remaining form of competition, and this would be a tragic result in an area so vital to the public health." [33] This is precisely the view of the subcommittee majority, although expressed from the point of view of an antipodal set of values. The Subcommittee was at great pains to prove that the "intensive competition for superiority" (*i.e.*, in profits) stimulated by patent availability is wasteful, duplicative, and supports high prices, but that the replacement of product differentiation by price competition would be tragic only to excessive profit levels, and would improve resource allocation in the drug market and eliminate the inequity of burdening the sick with high prices.

The drug industry tried to rebut these presumptions of noncompetitiveness by presenting a great deal of economic evidence regarding (1) price index data purporting to show a decrease in drug prices from 1949 to 1960, and (2) firm concentration and turnover in firm rank order of sales volume in various drug submarkets. Professor Markham of Princeton, retained by counsel for the PMA, presented a price index of prescription drugs, showing a decline of 7.6 percent for the period 1949–1960, or about one-half of one percent per year, compounded annually. This evidence, if valid, would argue for a remarkable degree of price stability in the industry, and might sustain the Subcommittee's loose usage of the loose term, "administered prices", one agreed-upon characteristic of which is that they do not typically increase during inflations to the extent that one would expect on the basis of short-run profit-maximizing theory. But what is the cause of this slight decline? Dr. Blair of the subcommittee staff pointed out that patented drugs almost never change in price, while non-patented drugs typically experience a substantial price decline over time, and suggested that the chain index might reflect the presence of a number of patented drugs with fixed prices, and a number of unpatented drugs which declined in price. He asked further information be submitted with regard to the relative price changes of patented and unpatented drugs separately, and also for other information, including a tabulation of the drugs and prices employed in the index. Apparently the computations on patented versus unpatented drug price trends were never submitted, although the other material requested was supplied. Anyone who takes the trouble to look through the basic price data used in computing the index will find that of the 329 drugs included, no fewer than 193, or 59 percent of the total, had experienced no price change during the period since their introduction in the index (not all drugs in the index were sold during 1949; many appear only in subsequent years); 95 drugs (28 percent of the total) had experienced price increases (*i.e.*, price in 1960 higher than price in initial year of appearance), and only 41 drugs (13 percent of the total) experienced price decreases.[34] This tends to confirm the hypothesis of price stability. It is of further interest to note that the BLS index of the prices of drug and pharmaceutical materials declined by 32.7 percent during this time, while its index for the cost of retail prescriptions increased by 27.8 percent.[35] It is perhaps evidence of lack of competition that prices of drugs dropped only 7.6 percent while prices of inputs dropped 32.7 percent; it is also curious that prescription costs increased 27.8 percent during this period if prescription drug costs actually fell by 7.6 percent, in view of the facts that (1) most prescription drugs are priced by druggists by adding a fixed markup (66⅔ percent) over wholesale price, and that (2) if anything, the labor and other compounding costs of prescriptions should have declined since 1949 as a result of the marked decrease in the percentage of prescriptions which requird compounding by the pharmacist, rather than the transfer of capsules to different bottles.

The concentration and turnover ratio data presented by Markham do indicate very considerable turnover among drug firms in relative sales rank. The conclusions reached as regards concentration as a whole must be interpreted in the light not only of the usual imperfections of concentration ratio data (especially the implicit assumption of independence of action among firms) but also the fact that the data refer to sales of the finished product rather than to production of the bulk powder, and there are great differences between concentration of these two in drugs. In 1958, for example, of the 51 most important prescription drugs, 27 were made by only one firm, 18 by two or three firms, and six by four to

[33] Ibid., pt. 4, p. 2003.
[34] *Ibid.*, pt. 4, pp. 2576–d through 2576–k.
[35] *Ibid.*, pt. 4, pp. 2469–2470.

seven firms; with regard to sales, only 17 were sold by only one firm, 12 by two or three firms, and 22 by four to thirteen firms.[36] Production of patented drugs is usually completely monopolized by the patent holder (unless there has been an interference settlement, with mutual licensing among contestants), but the sale of such drugs may not be monopolized. The patent holder may not have a sufficiently developed marketing network, enough detailmen, etc., to fully exploit the production monopoly, and may hence license other firms with wider marketing facilities to sell the finished product, such licensees tableting and bottling the bulk powder which the patent holder sells to them. For example, Carter licenses meprobamate ("Miltown") to Wyeth in order to increase its total sales.[37] While it is illegal for patent holders to dictate the pricing policies of licensees, price competition between such parties is almost unknown. But it is quite consistent with continual pure monopoly control of drug production that several of the licensees of the patent holder may alter their relative positions in the market from year to year. The turnover data presented by Markham do not indicate the extent to which this may be the case. The presumption made at the hearings that most of the turnover is induced by firms developing new products remains only a presumption; Mr. Mannis of Arthur D. Little, a participant in Markham's study, wasn't quite sure.[38] Markham placed great emphasis on the amount of turnover, but to assess the degree of workability of competition evidenced by such turnover, it is necessary to determine whether it was brought about by price competition, by genuine product competition, or by mere product differentiation with its accompanying wastes in the drug industry. The evidence (admittedly uneven in quality) of some ten thousand pages of Senate hearings indicates that a very small weight be placed on the first factor, a quite moderate weight on the second factor (less important since 1955), and the preponderance of weight on the last. But Markham seems to believe that turnover in itself is a good thing, and did not attempt to analyze its causes.[39]

According to Markham, the purpose of his study is "a thorough and objective appraisal of all economic aspects" of the drug industry.[40] But the study was sponsored by PMA, and PMA had its schedule. Markham continued, "We have begun by organizing our research schedule so as to center attention on the specific issues raised by the Subcommittee's report no. 448, entitled 'Administered Prices—Drugs,' published on June 27." (Markham's study began in July.) At the session of hearings when PMA presented its case in December, 1961, the study was apparently not yet complete. One of the issues slighted was price competition, although this issue had certainly been raised in the *Report*. During the examination on his testimony, Markham conceded that he had not examined the question as to whether or not any drug firm had experienced a change in relative sales rank because of price competition,[41] and Mannis also confessed to ignoring the role of price competition in effecting changes in concentration.[42] Markham agreed that price competition is of paramount interest to the consumer, but concluded his contribution to the Hearings with this statement: "I have not made any careful study of the workability of competition in the ethical drugs industry. I was examining primarily these particular issues that seemed to be important."[43] While a careful study of the workability of competition would admittedly take longer than the five months Markham had been able to devote to it prior to his

---

[36] Study of administered prices, *op. cit.*, p. 67.
[37] Ibid., pp. 17–18.
[38] Mr. Mannis testified, "Although I don't know, I would guess that back in 1951 or perhaps in the 1953 era, the leading product was quite different from the product that is leading the field now." *Hearings, op. cit.*, pt. 4, p. 2102.
[39] Ibid., pt. 4, p. 2106.
[40] *Ibid.*, Part 4, p. 2088.
[41] *Ibid.*, Part 4, p. 2096. Other attempts to support the existence of price competition were also less than convincing. Professor E. V. Rostow of Yale University, testifying in behalf of PMA, was asked to give instances where patented drug prices had weakened. His "best example," upon closer inspection, pertained to an instance where prices had temporarily declined, not because of competition among different patent-licensed manufacturers, but as a result of manufacturers disciplining their own wholesalers (who had sold at low prices to hospital formularies and in other competitive bid markets) by undercutting their prices, then cutting off their supplies, and later restoring the original prices. *Ibid.*, Part 4, pp. 2081–2082.
[42] Mr. Mannis attributed such changes as had occurred to "competition between and among products based on their merit. *Ibid.*, Part 4, p. 2102. This would be more acceptable if genuine product competition—even at identical prices—were prevalent in drugs; but when product differentiation becomes the road to sales, "merit" becomes what is measured by a popularity contest.
[43] Ibid., pt. 4, p. 2111.

appearance, it is quite another matter to seem to suggest that the workability of competition is not an important issue. On the contrary, the chief purpose of S. 1552 was to make the drug industry more workably competitive; it was not designed, for example, to stimulate or retard turnover in sales rank as an end in itself, or otherwise. But the point need not be labored.

B. *Senate Judiciary Committe Action*

S. 1552 was reported out of Senator Eastland's Judiciary Committee as a much more innocuous document than when it was submitted.[44a] All the patent provisions had been rejected. The other contemplated reforms were considerably attenuated. The request for public control over generic names on the part of the Secretary of the Department of Health, Education, and Welfare survived in modified form: under certain circumstances, the Secretary is to hold hearings and establish generic names for certain drugs. The request that the FDA pass on the efficacy of drugs became a provision that drug efficacy claims be supported by substantial evidence, but the wording adopted limited its jurisdiction only to initial claims made subequent to approval. The licensing requirement became mere registration, but the FDA was granted greater power to inspect plants. The requirement that advertisements contain generic names and statements on efficacy and side effects was entirely rejected.

Kefauver and other members of the Subcommittee majority noted that the amended bill would perhaps improve the quality of drugs, but would have little if any effect on prices, and indicated their intention to propose several amendments to the bill in order to salvage some of its economic impact. First, they proposed that all drug patent license agreements be filed with the Commissioner of Patents. It will be recalled that PMA had expressed willingness to make private patent interference settlements a part of the public record; the amendment would not require public filing, but would extend to all patent licensing agreements. Second, it was proposed that the compulsory patent licensing feature of the original bill be retained, but that it be invoked only where wholesale drug prices exceeded 500 per cent of factory costs. Third, it was pointed out that the wording of the current bill would not restrict drug claims to demonstrated efficacy after the approval of a new drug application, and requested the necessary modification in wording. Fourth, a compromise on the contents of drug advertisement was put forth: generic names need appear in type face only half as large as brand names, and summaries, rather than complete statements, of efficacy and side effects should be included.[44]

C. Congressional action

On August 3, 1962, the late President Kennedy communicated to Senator Eastland a list of amendents for the Committee to consider, relating chiefly to drug safety. The Committee met a few days later, and in addition to adopting some of the amendments suggested by the President, accepted Kefauver's amendments regarding advertising content and efficacy claims, but rejected his patent amendments. The bill was then transmitted to the House of Representatives, where further amendments were proposed. A conference report of October 3, 1962 shows that the conference between the Senate and House managers of the bill resulted in the substantial adoption of the Senate version, with only minor changes, except that the requirement that the Secretary of HEW publish and distribute to physicians all the detailed information required by law to be included in the drug package was dropped.[45] No reason is given for omitting this important provision. In the absence of information regarding motive or influence, the deprivation of easy access by the physician to the best information on drugs can only be interpreted as an act of disinterest misanthropy.

V. CONCLUSIONS AND POLICY RECOMMENDATIONS

The amended bill was passed by both houses and became law on October 10, 1962, as the "Drug Amendments of 1962," more frequently referred to as the Kefauver-Harris Act. It will serve to improve the quality of drugs, but its impact on competition will be slight. Requiring evidence of efficacy will, if properly en-

[44a] Rt. on S. 1552, op. cit., pp. 1–8.
[44] Ibid., pp. 36–48.
[45] Drug Amendments of 1962: Conference Rept. No. 2526, House of Representatives, 87th Cong., second sess., Washington, D.C., Government Printing Office, Oct. 3, 1962 pp. 10–18.

forced, keep off the market a certain number of useless but harmless drugs which presently encumber it. To the extent that drug firms are tempted to patent and push such inefficacious drugs in order to take advantage of the susceptibility of physicians to massive advertising, refusal to certify them will prevent further market confusion and reduce the amount of research devoted solely to pro viding vehicles for sales promotion.[46] But most research is probably not aimed at pure sales promotion of useless compounds; far more resources are wasted in duplicating existing drugs via molecular manipulation, and the failure of this part of S. 1552 to be enacted keeps incentives for such "research" as great as ever.

Competition between generic and brand name products usually exists only where the drug is not patented, *i.e.*, in only a small part of the prescription drugs market. The ability of generic drug firms to compete with brand name firms is slightly increased by the provisions for registration and inspection of plants. This should give physicians more confidence in the quality of generic name products and increase their willingness to prescribe by generic name, but in view of the demonstrated ability of drug firm detailment to disparage generic producers, it is doubtful if the fact of more adequate FDA inspection, in itself, will eliminate the cultivated distrust of generic drugs on the part of physicians. In the past, detailmen have also disparaged the scope and the adequacy of FDA inspection, and it remains to be seen if they will have similar success even under the new program. The only sure way of eliminating disparagement is to eliminate the army of over 15,000 detailmen; eliminating monopoly profits will largely bring this about. The provisions giving the Secretary of HEW authority to establish generic names supplies the basis for sweeping reforms resulting in shorter and simpler names physicians can remember, spell, and hence can prescribe. But much depends upon the aggressiveness with which the Secretary carries out this mandate. At present, few generic names do not need reforming, but a universal housecleaning in nomenclature is probably not to be expected. Yet it would be most disappointing if the Secretary were only to exercise his power in those relatively few cases where no name exists, or more than one.

The remaining reforms of advertising also proceed in the right direction, but not far enough. The requirements for summaries of side effects and efficacy in all advertisements will make for a more enlightened practice of medicine, but will not have direct effects on drug economics. The requirement that generic names be included in all copy will help, but the more such names are simplified, the more it will help.

To an economist concerned with questions of public policy, the congressional performance in regard to S. 1552 is discouraging, but not entirely so. The Act effected significant medical, if not economic, reforms. It now seems appropriate that renewed efforts be applied on the part of the public to induce congress to pass a more adequate act abolishing drug product patents and requiring generic name labeling. But the obstacles should not be underestimated. The influence of opposing groups is considerable, and the performance of several of these at the Hearings left much to be desired from the standpoint of concern for the public interest in health. Congress was apparently more sensitive to the need of protecting the consumer's physical health than his economic well-being; in fact the prime reason that any action was taken on the bill may have been its coincidence in time with the Thalidomide tragedy in late 1961, with the great publicity given the embryo-deforming effects of a new sedative drug. But it would seem preferable for the public to conduct a continuing campaign for more effective drug industry legislation, rather than to have such reform legislation come about piecemeal as a by-product of a series of tragic and avoidable blunders on the part of pharmaceuticals makers.

---

[46] Hearings on administered prices, op. cit., pt. 18, p. k0379 contains interesting testimony by Dr. A. D. Console, former medical director of Squibb, who in reply to Senator Kefauver's question as to how much drug firm research produces nothing worthwhile and is not intended to, answered: "I think the majority of it is in that category . . . with many of these products, it is clear while they are on the drawing board that they promise no utility; they promise sales."

Income Opportunities and Physician Location Trends in the United States

BY

HENRY B. STEELE and GASTON V. RIMLINGER

Reprinted from THE WESTERN ECONOMIC JOURNAL

Vol. III, No. 2, Spring, 1965

INCOME OPPORTUNITIES AND PHYSICIAN LOCATION TRENDS IN THE UNITED STATES*

HENRY B. STEELE and GASTON V. RIMLINGER
Rice University

INTRODUCTION

In a highly industrialized society professional manpower is one of the most important resources, which implies that trends in its geographic location have important economic and social implications. The spatial distribution of physicians in relation to population has been highly uneven in the United States for at least several decades.[1] Since this situation must be considered less than ideal from both the economic and the social points of view, the trends in the location of physicians should be of particular interest. The present article seeks to answer three basic questions. First, is there any tendency toward change in the distribution of physicians, and if so, is the long-run tendency toward a more even or a more uneven distribution pattern? Second, what is the role of physician location trends in the distributive pattern? Third, to what extent are the observed changes in physician location accounted for by such variables as population and per capita income, which may be taken to reflect income opportunities for physicians? It should be borne in mind that a more even distribution of physicians in relation to population can come about either because more physicians move to areas with a physician shortage or because more people move to areas with a relative physician surplus.

The time period covered by the analysis is from 1950 to 1959. This is a rather short period, but its choice is dictated by the fact that adequate and comparable data on physician location are available only for these two dates. The data are from the health manpower surveys of the U.S. Public Health Service.[2] The location trends over this period will be related to the following variables: (1) the degree of urbanization, (2) the regional shift in population, (3) the regional per capita disposable income at the beginning of the period, and (4) the increase in regional per capita income during the period. The choice of these variables is again limited partly by the availability of data and partly by the nature of the analysis, which is primarily statistical. A priori, the most important omission among variables would seem to be the income of physicians. Relevant comparative regional

* We are indebted to Mrs. Gloria Shatto for assistance in statistical computation. Responsibility for the article is solely that of the authors.

[1] An analysis of this distribution can be found in G. V. Rimlinger and H. B. Steele, "An Economic Interpretation of the Spatial Distribution of Physicians in the U.S.," *Southern Economic Journal*, July 1963, pp. 1–12.

[2] The 1950 data are from M. Y. Pennell and M. E. Altenderfer, *Health Manpower Source Book*, section 4, U.S. Public Health Service (Washington, D.C.: U.S. Government Printing Office, 1954); and the 1959 data are from W. H. Steward and M. Y. Pennell, *Health Manpower Source Book*, section 10, U.S. Public Health Service (Washington, D.C.: U.S. Government Printing Office, 1960).

data on physician income are available only for the year 1949.[3] An attempt will be made to investigate the significance of the physician income level at the beginning of our period in a separate comparison of standard metropolitan areas.

Trends in Physician Distribution and Location

A brief introductory comment on the meaning of the physician distribution may be in order. Concern with physician distribution is based on the problem of access to medical services. The number of physicians in relation to population in a given area may be used as an index of the availability of physician services, but it should be noted that this is only an approximation. Availability depends on the relationship of the supply and demand for services. The demand is determined not only by the size of the population, but also by such interrelated factors as the level and distribution of income, the age structure of the population, the possession of health insurance, the level of education, and the degree of urbanization. The supply of physician services depends similarly not only on the number of physicians but on the amount of time they devote to patients and their effectiveness in treating them. An area with a high concentration of retired physicians or with many physicians engaged in research or teaching naturally has a smaller supply of services for patients than a similar area where all physicians are engaged in active private practice. Unfortunately, we lack the data necessary to take these factors into account. The present analysis is based on the total number of physicians in an area rather than on the amount of service they provide. The areas used are for the most part groups of similar counties within a state; in one case standard metropolitan areas will be used. In each state counties are combined into the following groups adopted by the health manpower surveys: greater metropolitan, lesser metropolitan, adjacent, isolated semi-rural, and isolated rural.[4] Although some information regarding income differentials is lost by aggregating the 3069 counties of the United States into a little over 200 groups, the group is considered a preferable area concept as far as the market for medical services is concerned.

We turn now to the question of whether there was a change in the distribution of physicians between 1950 and 1959 and the direction of any such change. No change in distribution would mean that every area experiences the same relative increase or decrease in the physician-population ratio. In other words, if the national physician-population ratio increases by 5 per cent, an increase of 5 per cent of this ratio in every area of the country would leave the over-all distribution unchanged. Given the constant internal migration of the population and the entry and exit of physicians into and from professional life, it would be a very unusual coincidence if there were no change in distribution over a nine-year period. Hence, the question of whether there was a change in distribution is by itself not very meaningful. It is the extent and the direction of change that are

[3] These data are used below in the analysis of 19 standard metropolitan areas. For the source of the physician income data, see W. Weinfeld, "Income of Physicians, 1929–1949," *Survey of Current Business*, July 1951, pp. 9–26.

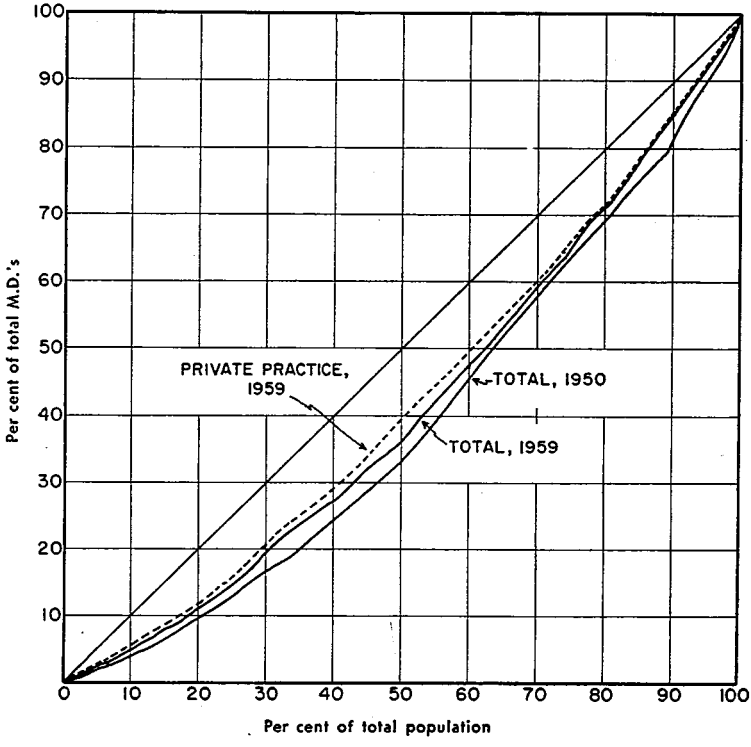
[4] For a precise definition of each of these, see Pennell and Altenderfer, *op. cit.*, section 4, pp. 18–20.

important. Physicians are most highly concentrated in high-income and in urban areas. After experimenting with several approaches, it developed that the most informative way to put the question of change in distribution is to inquire whether the degree of concentration of physicians in these areas has increased or decreased. The answer depends on how concentration is measured. It is interesting, and somewhat unexpected, that relative concentration in higher income areas decreased between 1950 and 1959, while at the same time the physician-population ratios in more highly urbanized areas increased. To clarify the meaning of these findings a more detailed examination is necessary.

To measure the change in relative concentration with respect to regional income, the writers constructed two Lorenz curves relating physicians to population, one for 1950 and one for 1959, as shown in Chart I. On the horizontal axis

CHART I

LORENZ CURVES FOR DISTRIBUTION OF PHYSICIANS BY RANK ORDER OF COUNTY GROUP PER CAPITA DISPOSABLE INCOME[a]



[a] Total physicians, 1950 and 1959; physicians in active private practice, 1959.

the population of the United States is arrayed in rank order of county group per capita disposable income,[5] with incomes rising from left to right. The vertical axis represents the total number of physicians in the country. It will be noted that the Lorenz curve for all physicians in 1959 lies closer to the diagonal than the corresponding curve for 1950, indicating a change toward a more even distribution of physicians with regard to regional income. It is interesting to note also that when only physicians in private practice are considered, the distribution in 1959 is more even than when all physicians are taken into account. Unfortunately, we do not have the necessary data to analyze the change in the distribution of private physicians between 1950 and 1959.

TABLE 1

CHANGE IN PHYSICIAN CONCENTRATION IN RELATION TO PER CAPITA REGIONAL INCOME, 1950–1959

Population ranked by income	*Per cent of physicians in 1950*	*Per cent of physicians in 1959*	*Per cent change 1950–1959*
Lowest 10%	4.0	4.8	+19.8
10–20	5.4	6.1	+13.5
20–30	6.7	8.4	+24.5
30–40	8.0	8.1	+ 0.9
40–50	8.9	9.6	+ 7.8
50–60	12.0	10.4	-13.3
60–70	12.6	11.8	- 6.1
70–80	11.6	12.2	+ 5.3
80–90	13.2	13.9	+ 5.1
90–100	17.5	14.9	-15.1

A clearer picture of the change in concentration can be obtained by examining Table 1, which shows the gains and losses of the regionally stratified income groups. On the whole the change was clearly in favor of lower income areas. The top 10 per cent of the income-ranked population suffered the greatest relative loss of physicians. The population with the greatest gain was in the third decile from the bottom of the income range. Of course, physicians still remain concentrated in high-income areas. In 1950 the bottom 25 per cent of the population had 13.0 per cent of the physicians and the top 25 per cent had 36.5 per cent of the physicians. In 1959 this had improved to 14.4 per cent for the bottom quartile

[5] Data on population and per capita disposable income are from *Sales Management Annual Survey of Buying Power*, May 1951, and May 1960. The income for the county group is a weighted average. The *Sales Management* staff terms its income measure "effective-buying income," but observes the same definition the U.S. Department of Commerce employs in defining disposable income.

and 35.5 per cent for the top quartile. The top 10 per cent of the population still had three times as many physicians in 1959 as the bottom 10 per cent, but in 1950 the top 10 per cent of the population had over four times as many physicians as the lowest 10 per cent. This improvement in the physician-population ratios of the lower income relative to the higher income areas should not be interpreted to mean that physicians relocated from high to low income regions. A more likely explanation is the tendency of the population to move to higher income areas, and the changes in regional income levels.

A valid argument can be made that for a general evaluation of the distribution, the concept of relative distribution developed above is more meaningful than the measurement of distribution based on degree of urbanization. Relative distribution takes into account the proportion of the population that gains or loses by a change in distribution. On the other hand an analysis in terms of the degree of urbanization shows what happens to specific social segments of the population and gives us another way of looking at the problem of distribution. It also has the advantage of isolating one of the important causal factors in the location of physicians and the relocation of the population as a whole. The results of this approach for the period 1950–1959 are summarized in Table 2.

An important fact which can be noted in Table 2 is that in none of the areas did the number of physicians increase as much as population or decrease as little as population. This means that the over-all decrease in the country's physician-population ratio between 1950 and 1959 left all aggregated county groups with fewer doctors in relation to population. However, not all areas shared equally in this relative loss of physicians. There was a redistribution in favor of the more highly urbanized areas and against the rural areas. In isolated rural areas population decreased slightly (.7 per cent), but the number of physicians decreased about six times as fast (4.4 per cent). This is the meaning of the 5.9 coefficient of elasticity of physician mobility.[6] This coefficient merely describes a numerical relationship and is not intended to indicate a causal significance of the same magnitude. Isolated semi-rural areas fared even worse than isolated rural areas, even though they had a slight increase (1 per cent) in the absolute number of physicians. Their problem is that population increased 11 times faster than the number of physicians. The other three county groups had somewhat comparable changes in physician-population ratios. Most favored were the greater metropolitan areas, where the relative increase in the number of physicians was equal to 70 per cent of the increase in population, as indicated by the elasticity coefficient of .7. In the lesser metropolitan areas, the relative increase in the number of physicians was second largest in the country, but the relative increase in population was the largest. As a result the lesser metropolitan areas had a rate of physician increase that was about half as fast as the rate of population increase. In the adjacent areas, where the rate of physician increase was much smaller, the rate of population increase was relatively still smaller. The result was a rate of physician increase equal to 60 per cent of the population increase.

[6] Computed before rounding the percentage changes.

TABLE 2

CHANGE IN PHYSICIAN DISTRIBUTION IN RELATION TO DEGREE OF URBANIZATION, 1950–1959

County group	*Per cent change in physicians (1)*	*Per cent change in population (2)*	*Average population elasticity of physician mobility (1) ÷ (2) (3)*	*Per capita income in 1950 (4)*	*Per capita income in 1959 (5)*	*Per cent increase in income (6)*	*Average per capita income elasticity of physician mobility (1) ÷ (6) (7)*
Isolated rural	−4.4	−.7	5.9	$ 849	$1,468	72.9	−.06
Isolated semi-rural	1.0	11.3	.1	1,046	1,506	44.0	.02
Adjacent	9.8	16.4	.6	1,287	1,975	53.5	.18
Lesser metropolitan ...	15.6	32.1	.5	1,439	1,888	31.2	.50
Greater metropolitan ...	18.7	27.9	.7	1,623	2,185	34.6	.54

Table 2 shows also some relationships between changes in physician-population ratios and regional per capita incomes. Column (7) shows the average per capita disposable income elasticity of physician mobility and reveals a somewhat surprising negative coefficient for isolated rural areas. A 1 per cent increase in per capita disposable income in an isolated rural region would be associated with a decrease of .06 per cent in the number of physicians. With increasing degrees of urbanization, the income elasticity increases: for isolated semi-rural areas, an increase in per capita disposable income of 2 per cent will increase the number of physicians by .02 per cent; the ratio increases to .18 per cent, .50 per cent, and .54 per cent for adjacent, lesser metropolitan, and greater metropolitan county groups.

There is a clear tendency toward higher income areas having smaller declines in the number of physicians in relation to population. This would seem to contradict the earlier finding about a change in relative distribution in favor of lower income areas. However, there is no necessary inconsistency, for two quite different ways of measuring changes in distribution are involved. Nevertheless, the apparent inconsistency does highlight the difficulty of describing changes in distribution in an unequivocal manner. Some questions may be raised also with respect to the relevant time period for which income is measured. This is especially the case when there is a change in the income rank over time. For instance, in 1950 lesser metropolitan areas had the second highest average income of the five county groups, but in 1959 they were in third place, behind adjacent areas. If the causal effect of income operates with a lag, the initial income level is more likely to be relevant. If it operates through anticipation, the final income level might be more appropriate, or even the rate of increase in income over time. Multiple regression and multiple correlation analysis is necessary to determine the respective significance of these different factors. This is the task of the following section.

Analysis of Physician Location Trends

The preceding discussion has shown that changes in physician location over time are an important aspect of the regional redistribution of physicians relative to population. A complete analysis of changes in distribution would have to account for the movements of the population as well as the movements of physicians. In this paper, however, we are concerned only with the movement of physicians. Population shifts will be taken as an independent variable that has a causal effect on the location of physicians. One would expect that an increase in population in a given area, especially if it is through migration, makes it easier for physicians to open new practices and induces them to locate there. Of course, once a doctor has established his practice in a given area and built up his clientele, he is not likely to move. The highly personal nature of the doctor-patient relationship is undoubtedly an important barrier to mobility. On the other hand, new doctors are constantly entering the profession while others are retiring and leaving vacancies. It is mainly through these entries and exits that physicians are relocated, which makes it important to add a substantial time dimension to the

analysis of their mobility. It is the mobility of the group rather than that of the individual that counts.

In addition to increases in population, Table 2 suggests that the regional level of income may act as a strong inducement to new physicians. The simple prospect of a wealthy clientele might have much drawing power. People in the higher income brackets might be in the habit of paying their medical bills more promptly than persons with low incomes. But if the drawing power is based primarily on the physician's expectation to be able to earn a high income by charging higher fees, it will be greatly dampened if the high-income area already has a very large number of doctors in relation to population. On the other hand, a high concentration of doctors (provided that it does not affect income too adversely) might be attractive because it offers more opportunity for leisure. It is reasonable to assume that where there are many doctors it is easier to find vacation or weekend replacements than where there is a shortage.

The results of an attempt to evaluate the relationship of changes in the regional number of physicians to changes in a selected set of variables are presented in Tables 3 and 4, which show the coefficients of regression and of correla-

TABLE 3

REGRESSION COEFFICIENTS PERTAINING TO PERCENTAGE CHANGES IN THE NUMBER OF ALL PHYSICIANS IN SIX SECTORS, 1950–1959[a]

County group	*Regression constant (a)*	*Coefficient of per cent change in population (b)*	*Coefficient of per capita income in 1950 (c)*	*Coefficient of per cent change in income (d)*
Isolated rural	−40.43	.64	.02	.03
Isolated semi-rural	−29.34	−.22	.40	.05
Adjacent	48.26	1.42	−.50	.14
Lesser metropolitan	56.72	1.00	−.08	.29
Greater metropolitan	48.86	.92	−.01	.61
			Coefficient of physician income in 1949	
Nineteen standard metropolitan areas	50.33	.96	.01	.00[b]

[a] The value of the coefficients is derived from the linear regression equation:

$$\frac{\Delta N}{N_{t(o)}} = a + \frac{b\Delta P}{P_{t(o)}} + c I_{t(o)} + \frac{d\Delta I}{I_{t(o)}},$$ where N represents the number of physicians; I, the regional per capita income; P, population; and $t(o)$, the initial year of the period.

[b] Less than .005.

TABLE 4

CORRELATION COEFFICIENTS PERTAINING TO PERCENTAGE CHANGES IN THE NUMBER OF ALL PHYSICIANS IN SIX SECTORS, 1950–1959

County groups	TOTAL CORRELATION	PARTIAL CORRELATION			SIMPLE CORRELATION		
	All variables ($R_{1.234}$)	*Population and income level* ($R_{1.23}$)	*Population and income change* ($R_{1.24}$)	*Income level and income change* ($R_{1.34}$)	*Population* ($R_{1.2}$)	*Income level* ($R_{1.3}$)	*Income change* ($R_{1.4}$)
Isolated rural	.33	.16	.20	.14[a]	.14	.06[a]	.12[a]
Isolated semi-rural	.63	.61	.59	.28	−.57	.26	.21
Adjacent	.87	.82	.82	.08[a]	.82	−.01[a]	.06[a]
Lesser metropolitan	.93	.90	.92	.88	.87	−.32	.82
Greater metropolitan	.67	.48	.63	.64	.44	−.37	.53
		Population and physician income		*Physician income and income change*		*Physician income*	
Nineteen standard metropolitan areas	.95	.95	.95	.17	.95	.17	.06[a]

[a] Not significant at the .05 level.

tion of these variables. The tables relate the average rate of change in the number of physicians in each of the county groups to (1) the average rate of population change, (2) the average level of absolute income in 1950, and (3) the rate of change of the average per capita income during 1950–1959. These national averages are simple arithmetic averages of state data, which in turn are weighted averages of county data within each county group. In the case of the 19 standard metropolitan areas, the averages are simple arithmetic averages, but the level of income of the population in 1959 is replaced by the average level of physician income in 1949, which adds a fourth variable to the analysis. The fifth variable is the degree of urbanization, which is treated as a qualitative variable. It could be included in the regression and correlation equations by means of dummy variables, but we chose the alternative of computing separate regressions and correlations for each county group [7] and for 19 standard metropolitan areas. This has the effect in each case of keeping the degree of urbanization constant and provides a basis for comparison between urban and rural areas.

One of the obvious facts indicated by Table 3 is that the impact of the population and income variables depends greatly on the degree of urbanization. It is well known that the rural environment tends to discourage and the urban environment to encourage the location of physicians. There are many factors, especially differences in cultural and professional advantages, which account for this differential in drawing power. But here we are concerned only with variables that affect the income opportunities of physicians. To what extent do they have a differential impact in the different environments? We might like to know whether improved income opportunities would work in attracting more physicians to rural areas and roughly how much of an improvement would be required. Of course, our data can only give crude indications of what the answer to these questions might be.

Let us compare now the effects of our variables in the different environments. Column (b) of Table 3 shows the percentage increase in the number of physicians associated with a 1 per cent increase in population in the different county groups. On the whole the effect of population change is quite strong, but it varies greatly between areas. It is greatest in the adjacent areas, where a 1 per cent increase in population is associated with a 1.42 per cent increase in the number of physicians. Somewhat puzzling is the case of semi-rural areas, where a 1 per cent increase in population is associated with a .22 per cent decrease in the number of physicians. In lesser metropolitan areas, population and physicians tend to increase at the same rate, while in more urbanized areas the rate of physician increase is somewhat smaller than the rate of population increase, and in isolated rural areas it is much smaller. The basic distinction is clearly between rural and nonrural areas. Assuming the observed association to be of a causal nature, we could say that

[7] A few counties changed from one group to another over the nine-year period in question, usually moving from lesser to greater urbanization categories. In these cases the county was included in the category which it occupied in 1950, on the assumption that its condition in the initial year had a greater impact than its classification at the end of the period. The total effect of such changes, however, is negligible since only 34 out of 3069 counties were involved.

in rural areas population increases tend to lower the physician-population ratio substantially, whereas in metropolitan areas these ratios would be little affected and in adjacent areas they would increase.

The extent to which the average income level in 1950 is associated with the increase in the number of physicians is shown in column (c) of Table 3. The variations between areas are again very great and this time rather unsystematic. In isolated rural areas, a $100 higher per capita income in 1950 is associated with a 2 per cent increase in the number of physicians over the 1950–1959 period. On the other hand, in isolated semi-rural areas, a $100 higher income is associated with a 40 per cent increase in physicians, and in adjacent areas with a 50 per cent decrease. These erratic results raise questions about the significance of the observed association, which will be explored below. With respect to the 19 standard metropolitan areas, the observed relationship indicates a 1 per cent increase in the number of physicians over 1950–1959 for a $100 higher level of physician income in 1949.

Column (d) of Table 3 describes the associations between increases in physicians and increases in regional per capita income over 1950–1959. In the 19 standard metropolitan areas a systematic relationship is for practical purposes nonexistent. But it is interesting that in the other areas the association becomes constantly closer with higher degrees of urbanization. In isolated rural areas a 1 per cent increase in income is associated with only a .03 per cent increase in the number of physicians, whereas in greater metropolitan areas it is associated with a .61 per cent increase in physicians.

We have now measured the relationships between changes in selected variables and corresponding changes in the number of physicians. But it must be remembered that all these changes are averages and that their significance depends on the scatter of the original values about the regression line. An indication of this significance is presented in Table 4, which contains the correlation coefficients for the population and income variables. On the whole the correlations are highest in the 19 standard metropolitan areas and the lowest in the isolated rural areas. As a matter of fact, the variables considered in this study, other than the changes in population, are of little use in explaining changes in the number of physicians in isolated rural areas because the simple correlation coefficients do not appear to be statistically significant. In isolated semi-rural and in greater metropolitan areas the situation is slightly better. The index of total correlation $R_{1.234}$ for the semi-rural areas is .63, which makes the index of total determination $R^2_{1.234} = .40$. In other words, 40 per cent of the observed change in the number of physicians in the semi-rural areas over the nine years is "explained" by population change, the initial income level, and the change in income. In greater metropolitan areas 45 per cent of the change is explained by these variables. However, in adjacent, lesser metropolitan, and the 19 standard metropolitan areas, the corresponding percentages are 66, 86, and 90, respectively.

Among the variables examined, population change is the one most closely correlated with change in the number of physicians. Total correlation is generally

not greatly affected by dropping either the initial level of income or the change in income as independent variables. In the case of the adjacent, the lesser metropolitan, and the 19 standard metropolitan areas, total correlation is not greatly affected even if both of these variables are dropped. Only in lesser metropolitan areas is there a high simple correlation between change in income and change in physicians. In these same areas the combination of income level and income change also has the highest correlation with changes in the number of physicians. The level of physician income at the beginning of the period has only a very low correlation with the subsequent increase in physicians. Finally, it should be noted that, on the whole, the change in physicians is more closely correlated with changes in per capita income than with the initial level of per capita income. This is especially the case in the lesser and the greater metropolitan areas.

Summary and Conclusions

It has been shown that the two most important factors affecting the change in location of physicians over time are the regional degree of urbanization and the increase in population. The extent to which an area is urban or rural, or adjacent to an urban area, influences its attractiveness to physicians aside from what is happening to its population or to its level of income. An increase in population has an effective drawing power in urban but not in rural areas. In the isolated rural areas as a whole both population and physicians decreased, but there was no meaningful correlation between these two changes. Rather surprisingly, the level of per capita income at the beginning of the period turned out to have little effect on physician location. Nor were physicians induced to locate on the basis of higher average physician incomes. Increase in per capita income seems to have been a factor in attracting more physicians but only in lesser and greater metropolitan areas. The fact that the isolated rural areas had the highest rate of increase in per capita income did not prevent a decrease in the number of physicians. Even when initial per capita income and increase in income are combined, their effect is strong only in lesser and greater metropolitan areas. This indicates that on the whole the income of the population is a less significant factor than we had originally suspected.

However, in evaluating the meaning of these findings it is important to bear in mind that they refer to all physicians, rather than only to those engaged in private practice. The Lorenz curves in Chart I show that there is a substantial difference between the distribution of all physicians and that of private physicians. It is possible therefore that if a comparison over time were made of physicians in private practice, the income of the population would be a more important factor. At least this was suggested by our analysis of the distribution of private physicians in 1959.[8] It is quite clear that a more disaggregated analysis is necessary to obtain more precise and more reliable results. For instance, it might be worth while to break down the increase in population into increase through migration and natural increase. It might be profitable also to analyze changes in the number of physicians

[8] See Rimlinger and Steele, *op. cit.*

by examining both entries and exits in a given area. The age structure of the physicians would be important in this respect, since it affects both their income and their exit from the profession. Additional insight would be gained by studying separately the location trends of specialists and general practitioners. Table 5 shows that the degree of urbanization is a much stronger factor in the location of specialists than in the location of general practitioners.

TABLE 5

PRIVATE NON-HOSPITAL PHYSICIANS PER 100,000 POPULATION IN THE UNITED STATES, 1959

County group	*General practitioners*	*Full-time specialists*
Isolated rural	42.8	2.6
Isolated semi-rural	46.5	22.9
Adjacent	46.4	20.0
Lesser metropolitan	42.5	55.4
Greater metropolitan	51.6	64.4

SOURCE: *Health Manpower Source Book*, section 10, p. 12.

A final comment may be in order with regard to the general problem of physician distribution. It was shown that the relative inequality of distribution based on regional per capita income has decreased while the urban-rural inequality has increased. This is consistent with the findings of the physician location trends. Since the pull of population increase is stronger than the attraction of per capita income, a redistribution in favor of lower income areas is possible. The change in distribution, however, seems to be dominated by the movement of population rather than the movement of physicians. Population tends to move to higher income areas and physicians follow but at a slower rate. This has the long-run tendency of reducing the concentration of physicians in higher income areas, while at the same time it is not necessarily inconsistent with a continuation of great urban-rural inequalities. But it must be remembered that an increasingly smaller percentage of the total population will be in areas with low physician-population ratios. Complete elimination of higher physician-population ratios in higher income areas is not likely to occur, nor is it necessarily desirable from an economic point of view. How fast the equalization of distribution will proceed depends in part on how the population redistributes itself and how fast physicians follow the movement of the population. Obviously, if population increases in higher income areas did not attract more physicians, the relative concentration of physicians in these areas would be more quickly eliminated.

Reprinted from SOUTHERN ECONOMIC JOURNAL
Vol. XXX, No. 1, July, 1963
Printed in U.S.A.

AN ECONOMIC INTERPRETATION OF THE SPATIAL DISTRIBUTION OF PHYSICIANS IN THE U. S.*

GASTON V. RIMLINGER
AND
HENRY B. STEELE
Rice University

I. INTRODUCTION

The number of physicians in relation to population varies widely between localities and regions in the United States. This pattern of distribution is related to differences in regional per capita income levels and raises questions not only from a public health but also from an economic point of view. To what extent and in what manner do economic factors influence the distribution of physicians? What kind of economic behavior of physicians is suggested by their spatial distribution? Do the economic forces at work within the given institutional arrangement tend toward a distribution that is consistent with efficient use of health manpower resources?

The purpose of this article is to seek an answer to these questions.[1] It begins with a description of the empirical relationship between the ratios of physicians to population and regional per capita incomes. Several hypotheses are then developed to explore the economic variables that may account for the observed relationship. Finally, the hypotheses are subjected to empirical tests, insofar as is possible with the data available on physician incomes, physician visits, and medical expenditures. Although relevant data in this area are limited, the article does bring together in a concise form a large amount of statistical information from both public and private sources.

II. THE DISTRIBUTION OF PHYSICIANS

Any analysis of the spatial distribution of physicians poses the problem of defining the geographic area served by physicians. Following the practice of the Health Manpower Survey,[2] we shall assume that the "county group" is the relevant economic area. A county group combines counties with similar population concentrations within a state; these counties are often but not always geographically contiguous.[3] The

* We are indebted to Professor Edgar O. Edwards and Dr. George J. Benston for their helpful comments on an earlier draft and to Mr. Vincent Tarascio for his assistance in compiling statistical information. Responsibility for the article is solely our own.

[1] Similar questions may be raised with regard to distribution between specialties. None of these questions has been adequately explored by economists. See, however, Rashi Fein, "Studies on Physician Supply and Distribution," *American Journal of Public Health*, May 1954, pp. 615–624; and M. Friedman and S. Kuznets, *Income from Independent Professional Practice* (New York: National Bureau of Economic Research, 1954).

[2] The Health Manpower Surveys used in this study are in M. Y. Pennell and M. E. Altenderfer, *Health Manpower Source Book*, Section 4, U. S. Public Health Service (Washington, D. C.: U. S. Government Printing Office, 1954); M. Y. Pennell and M. E. Altenderfer, *Health Manpower Source Book*, Section 5, U. S. Public Health Service (Washington, D. C.: U. S. Government Printing Office, 1954); and in W. H. Stewart and M. Y. Pennell, *Health Manpower Source Book*, Section 10, U. S. Public Health Service (Washington, D. C.: U. S. Government Printing Office, 1960).

[3] Five groups are used and they are designated as follows: (1) greater metropolitan, where the populations exceed one million; (2) lesser metropolitan, where populations range from 50,000 to one million; (3) adjacent counties, those which are contiguous to metropolitan areas; (4) isolated semirural areas, for counties with at least one city of 2,500 people; and (5) isolated rural areas, for the remaining rural counties. Residents of adjacent areas are no doubt served on occasion by

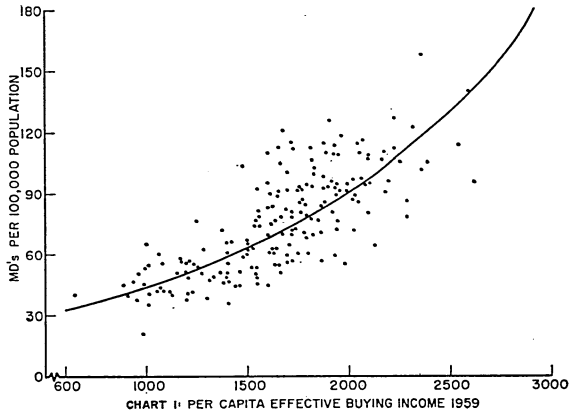


CHART I: PER CAPITA EFFECTIVE BUYING INCOME 1959

data presented in Chart I thus are based on the number of physicians, the size of the population, and the per capita income of the population in each county group for the entire country in 1959.[4] Each observation is for a county group, of which there are one to five in each state.

The distribution is stated in terms of privately employed physicians, both general practitioners and specialists, per one hundred thousand persons in a county group. This measure excludes all federal physicians, all retired or inactive physicians, and all full-time hospital, teaching, and research physicians. In an analysis of the role of economic variables in the spatial distribution of physicians these exclusions seem to be clearly desirable. The income measure to which the physician-population ratio is related in Chart I is the county group per capita net effective buying income.[5]

The most significant aspect of the distribution presented in Chart I is the tendency of the physician-population ratio to rise with increases in regional per capita income.[6] For instance, at the $1,000 per capita income level the average ratio is about 45 physicians per 100,000 population, but at the $2,000 level it is around 90 physicians per 100,000 population. This doubling in the number of physicians in relation to population is not due solely to changes in income. Low income county groups are almost exclusively rural while high income county groups are metropolitan areas. Income and environment change together, which makes it impossible to separate their impacts. There are valid reasons for believing that both higher incomes and urban environment attract more doctors, but we are concerned primarily with the question of how larger incomes may account for higher physician-population ratios.

III. THEORETICAL INTERPRETATION

This section develops hypotheses to explain the rise of the physician-population ratio with increases in per capita income.

metropolitan physicians, which tends to understate the effective physician-population ratio for adjacent areas and to overstate it for metropolitan areas. Isolated areas are more likely to be self-sufficient.

[4] The data for Chart I are from *Health Manpower Source Book*, Section 10.

[5] Net effective buying income is identical with "disposable personal income" as computed by the Department of Commerce. See *Sales Management*, May 1961, pp. 54–56. The physician-population ratios and the per capita incomes are weighted averages of all counties of a given county group.

[6] The relationship is mildly curvilinear as indicated by the equation $y = 21.24(2.058)^x$ for the regression line fitted by least squares.

A. Linear Model

Let us begin with the simple case in which it is assumed that all physicians are interested in maximizing money income. We may further assume either that all physicians are of equal ability, or that patients have no way of determining physician ability; in either event, there are no quasi-rents attributable to ability or reputation. It is also assumed that all physicians are in the habit of charging what the traffic will bear, and that this amount is a fixed per cent of the patient's income.[7] Finally, it is assumed that there is no price competition among physicians and that the number of patient visits per physician is a function of the physician-population ratio alone, which rules out any relationship between per capita income and the number of patient visits per time period.

On the basis of this set of highly simplified assumptions we shall set up a theoretical model using the following symbols:

a = proportion of income spent on the average visit, which is constant
b = number of visits per person per unit of time, which is also constant
Y = national per capita income
Y_r = regional per capita income
N = total number of physicians
P = population
R = P/N, national physician-population ratio
R_r = regional physician-population ratio
y = income of physicians, which in equilibrium is the same for the entire country

The equilibrium income of physicians is equal to the aggregate expenditure on physician services divided by the total number of physicians.

$$(1) \qquad y = \frac{(abY)P}{N} = \frac{abY}{R}$$

In any given region we have

$$(2) \qquad y = \frac{abY_r}{R_r}$$

and

$$(3) \qquad R_r = \frac{ab}{y} Y_r$$

By thus determining the relationship between R_r and Y_r we have defined the spatial distribution of physicians in relation to population and regional per capita income.

This distribution pattern is illustrated in Figure 1 for various total numbers of physicians in the country. The number of physicians on the graph increases from N_0 to N_3 while the corresponding physician income decreases from y_0 to y_3. The physician-population ratio in any area is a linear function of regional per capita income and is represented by a straight line starting from the origin. The larger the total number of physicians (N), other things being equal, the smaller the average physician income (y) and the greater the slope (ab/y) of the line representing the distribution of physicians.

Several characteristics of this distribution may be noted. (1) If physicians and patients behave as the model assumes they do, the physician distribution will be a direct consequence of the regional distribution of income. (2) The absolute difference in the number

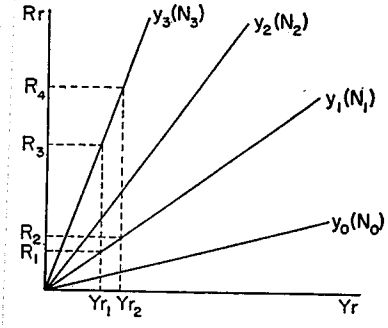


FIG. 1.—Linear Model of Physician Distribution. Regional Physician-Population Ratio (R_r) varies directly with Regional Per Capita Income (Y_r). The total number of Physicians (N) determines the level of physician income (y).

[7] The percentage will vary among different types of visits, but it is not necessary to introduce this complication.

of physicians in relation to population between any two areas depends on the difference in income, the proportion of income, a, spent on each visit, the number of visits per person per unit of time, b, and the total number of physicians, N. However, (3) the relative distribution of physicians, as shown by a Lorenz curve, is independent on the magnitude of either a, b, or N.

The effect on distribution of an increase in N can easily be demonstrated graphically. In Figure 1 an increase in the number of physicians from N_1 to N_2 raises the physician-population ratio in area Y_{r_1} from R_1 to R_3 and in area Y_{r_2} from R_2 to R_4. The absolute increase in the ratio is greater in area Y_{r_2} than in area Y_{r_1}, since $R_4 - R_2 > R_3 - R_1$, but the relative increase is the same in both areas, since $R_2/R_1 = R_4/R_3$. The percentage increase in the physician-population ratio in any area will be the same as the percentage increase in the total number of physicians.[8]

The distribution of physicians in this situation could therefore not be made more equal by simply increasing the number of physicians; it would in fact become less equal in an absolute sense. Nor would a willingness on the part of physicians to reduce their fees in relation to income result in a more equal relative distribution, unless fees were reduced to zero. A change in fees or in the number of visits per person would have the same effect on physician distribution as a change in the number of physicians, since they all affect the slope of the distribution line.

B. Nonlinear Models

It is not consistent with known facts to assume that the number of visits per person remains constant as per capita income increases. As a rule higher incomes are correlated with larger numbers of visits, although the absolute increase is not likely to be very large. We can easily take this relationship into account by expressing the number of visits as a function of income. Let

$b'Y$ = number of visits per person per unit of time in the country as a whole, and

$b'Y_r$ = regional number of visits per person per unit of time.

The regional physician income then becomes

$$y = \frac{ab'Y_r^2}{R_r} \tag{4}$$

and the regional physician-population ratio

$$R_r = \frac{ab'Y_r^2}{y} \tag{5}$$

The relationship between regional income and the physician-population ratio is now nonlinear, and the distribution lines are curves convex to the origin.

Other things remaining equal, the introduction of a positive relationship between visits and income makes the physician distribution more unequal, in the sense of pushing the Lorenz curve to the right. However, as before, an increase in the number of physicians would still leave the relative distribution unaffected and increase the physician-population ratio everywhere by the same percentage.[9]

[8] This can be demonstrated mathematically by showing that the elasticity of the regional physician-population ratio with respect to changes in the number of physicians equals one:

$$E_{R_r, N} = \frac{dR_r}{R_r} \Big/ \frac{dN}{N} = 1$$

By substituting $\frac{(abY)P}{N}$ for y, we get

$$R_r = \left(\frac{Y_r}{YP}\right) N \quad \text{and} \quad \frac{dR_r}{dN} = \frac{Y_r}{YP}$$

We can now show that:

$$\frac{dR_r}{R_r} \Big/ \frac{dN}{N} = \frac{Y_r}{YP} \cdot \frac{N}{R_r} = \frac{Y_r N}{YP} \cdot \frac{1}{Y_r N/YP} = 1$$

[9] As in the previous case, the elasticity of the physician-population ratio with respect to changes in number of physicians equals one. From equation (5) we get

$$R_r = \frac{Y_r^2}{Y^2} \cdot \frac{N}{P}$$

There may also be some question as to whether the fee remains a constant proportion of income as income changes. An increasing proportion would have the same effect as that described above for an increasing number of visits. A decreasing proportion, which is more likely, would tend to bend the curves in the opposite direction.

We can also get a nonlinear relationship between physician-population ratios and income, and hence a distribution that tends to be more unequal than that of our first model, if we abandon the assumption that physicians try to maximize only income. For the sake of simplicity, let us return to our initial assumptions of constant fees in relation to income and a fixed number of visits per person regardless of income. In equilibrium, all doctors have identical incomes, but those in low income areas work very long hours, as indicated by the low physician-population ratio, while those in high income areas have considerable leisure. Since there is freedom of movement between areas, it seems likely that some doctors will be tempted to leave the low per capita income areas and move to higher per capita income areas, where the same physician incomes are being earned with less work. This would result in a spatial redistribution of physicians and physician incomes. At the new equilibrium, each doctor would maximize the utility derived from both leisure and income. Some doctors may actually move from higher to lower income areas as the influx into high income areas lowers income there and increases leisure. But if doctors in general are willing to work harder only if they get a higher income,[10] there will be at the new equilibrium fewer doctors, earning higher incomes, in low per capita income areas and more doctors, earning lower incomes, in high per capita income areas. Somewhere in the intermediate per capita income range the physician-population ratio and the physician income would remain unchanged. The extent of the redistribution depends on the physicians' income-leisure preference pattern.

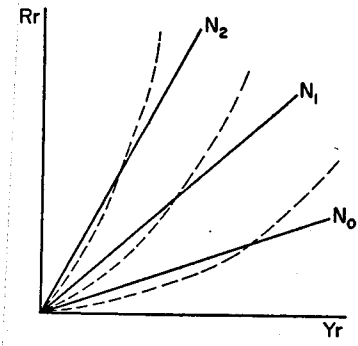


FIG. 2.—Linear and Nonlinear Models of Physician Distribution contrasted. Regional Physician-Population Ratio (R_r) increases non-linearly with increasing Regional Per Capita Income (Y_r), for various assumed total numbers of Physicians (N).

Figure 2 illustrates the impact of income-leisure substitutions for three different numbers of physicians in the country and initial levels of physician income. The convex curves represent the new equilibrium physician-population ratios in relation to per capita income of the population. In this case, the high income areas have increased the ratios and the low income areas have reduced them. How a change in the number of physicians, in fees, or in the number of visits would affect the distribution of physicians under these circumstances cannot be determined *a priori* without specifying the physicians' income-leisure preference pattern.

Hence

$$\frac{dR_r}{dN} = \frac{Y_r^2}{Y^2 P}$$

Thus

$$E_{R_r, N} = \frac{dR_r}{R_r} \Big/ \frac{dN}{N} = \frac{dR_r}{dN} \cdot \frac{N}{R_r}$$

$$= \frac{Y_r^2}{Y^2 P} \cdot \frac{1}{Y_r^2/Y^2 P} = 1$$

[10] Technically, their marginal rate of substitution of income for leisure is less than zero and leisure is not an inferior good.

C. Role of Mobility

The analysis thus far indicates that, as long as the fee rises with income and the use of physician services per person is fixed, physicians in low income areas will tend to have either less leisure or less income than physicians in high income areas, but not both. This follows in part from the assumption of mobility and equal ability. If these assumptions are relaxed the result no longer necessarily holds. Absence of mobility from low to high income areas may leave doctors in these areas with both less income and less leisure than in high income areas. On the other hand, a preference by doctors for high income areas has the same effect as a substitution of leisure for income; it tends to increase leisure and reduce income in high income areas and cause the opposite in low income areas.

The relaxation of the assumption of equal ability may also affect the results obtained. Specialists, who represent a group with higher than average qualifications, are likely to draw higher fees than general practitioners. Since they choose to locate mainly in high income metropolitan areas, they may raise both the leisure and the average physician income in these areas.

D. Summary and Conclusion

The main determinants of physician distribution, given the regional distribution of income, were found to be (1) the relation of fees to patient income, (2) the relation of demand for physician services to patient income, and (3) the behavior of physicians with respect to price competition, income maximization, desire for leisure, and geographic mobility. The degree of inequality of physician distribution tends to be greater the greater the increase in fees and in demand for services for a given increase in interregional income and the greater the inclination of physicians to substitute leisure for income.

An interesting insight provided by the theoretical analysis is that so long as expenditures on physician services rise with increases in income, and physicians do not engage in price competition but try to maximize income, changes in the total number of physicians, or in the general level of fees, or in the level of demand for services, do not affect the relative distribution of physicians. However, increases in the number of physicians, in level of fees, and in level of demand would increase the absolute physician-population ratio differentials between high and low income areas and in that sense would make the distribution more unequal.

IV. EMPIRICAL ANALYSIS OF DISTRIBUTION DETERMINANTS

We shall now examine to what extent available evidence on physician behavior, fees, and demand for services supports the preceding theoretical explanation of physician distribution.

A. Income Maximization, Leisure, and Mobility

In investigating the behavior pattern of physicians we shall make the assumption, which will be substantiated later, that medical expenditure per person (fees times visits) rises with income. This will tend to give us an uneven distribution of physicians in favor of high income areas, unless physicians ignore income. There is no evidence of such behavior. Nor are there data available that would allow a positive test of the income maximization hypothesis. The only feasible approach is to test the behavior pattern for its consistency with income maximization and to find out in this manner whether substitution of leisure for income and lack of mobility should be considered significant factors in the observed distribution of physicians. Lack of mobility is intended to cover all reasons for not moving other than income and leisure.

If leisure is valued by physicians, and if they are fairly mobile, we should find higher physician incomes where leisure is low and

lower incomes where leisure is high. This pattern seems to emerge in a test of thirty-two major cities.[11] The test, however, is not based on a direct measure of leisure but assumes that the higher the physician-population ratio the greater the amount of leisure available to physicians. For the cities tested, high physician incomes are correlated with low physician-population ratios (low leisure) and lower physician incomes with higher ratios (more leisure).

If, on the other hand, leisure does not increase with higher physician-population ratios, low incomes are not compensated by more leisure. This would most likely indicate a lack of concern with leisure and would make lack of mobility a significant factor in the distributive pattern. In this instance lack of mobility makes the distribution more uneven than it would have been under income maximization. If physicians tried to maximize income, there would be a redistribution from the low income (high ratio) areas in favor of the areas with higher incomes and fewer doctors.

An investigation of data on physician visits in eight major cities supports the hypothesis that lack of mobility rather than desire for more leisure is the major factor in the distribution of physicians among cities.[12] Physicians in those cities having the higher physician-population ratios do not appear to have more leisure than physicians in other cities. Hence, in the case of the 32 cities considered above, it was probably not a desire for greater leisure but some obstacle to mobility that kept physicians from moving to towns where there were fewer doctors.

A careful examination of the available data leaves a clear impression that desire for leisure is not a strong motivating force in the physicians' choice of location. Their low valuation of leisure is reflected in the fact that they tend not to reduce the length of their workweek as their income increases. The percentage of doctors working 49 hours or more per week is about the same at all income levels and ranges between 70 and 85.[13] It might be, of course, that physicians at higher income levels take longer and more frequent vacations.

When we compare physician incomes among communities of greatly differing sizes it becomes extremely difficult to deduce a rational behavior pattern. We find the lowest physician incomes in the smallest and largest size communities, those with under 2,500 and over 1,000,000 population. These are the areas where physician-population ratios tend to be the lowest and the highest respectively. We must invoke special kinds of immobility to account for the fact that physicians in very small communities tend to have less income and probably also less leisure than their colleagues elsewhere. Perhaps those in the profession who are least qualified tend to locate there. The high concentration of physicians in very large cities, despite lower incomes, may be partly explainable by cultural and professional

[11] This test uses 1949 income data for physicians and 1950 data on physician distribution. The regression on physician income (y) of the physician-population ratio (x) can be expressed by the equation: $y = 22.106 - .556\,x$, where y is expressed in terms of thousands of dollars per year and x is the number of physicians per hundred thousand population. The coefficient of correlation is .808, and is significant at the .05 level. Income data are from W. Weinfeld, "Income of Physicians, 1929–1949," *Survey of Current Business*, July 1951, pp. 9–26. Physician data are from *Health Manpower Source Book*, Section 4.

[12] The test is concerned only with physician visits outside of hospitals, for a period extending from July 1957 to June 1959. If we omit Boston (which is very unusual in that its ratio of physicians not in private practice to total physicians is by far the highest of the eight cities), the relationship between per capita visits (x) and the physician-population ratio (y) may be expressed by the regression equation: $y = -4.57 + 28.83x$. The coefficient of correlation is .963. Data on visits are from U. S. Public Health Service, *Health Statistics*, Series C, No. 6, p. 25. The high apparent correlation may be partly spurious in view of the presence of the population factor in the denominators of both the dependent and the independent variables. However, the statistical issue raised by this kind of ratio correlation remains unsettled at present. See J. R. Meyer and E. Kuh, "Correlation and Regression Estimates when the Data are Ratios," *Econometrica*, October 1955, pp. 400–416.

[13] *Health Manpower Source Book*, Section 5, pp. 79–81.

advantages and by opportunities to practice certain specialties not available elsewhere, which constitute another form of immobility.

In our estimation, improved mobility, in the direction indicated by income maximization, would result in a more even distribution of physicians among large cities, but might very well result in a more uneven distribution between large cities and smaller communities. Whether either redistribution would be desirable in terms of improving the relative availability of medical services depends, of course, on the structure of the demand for services.

B. Demand for Physician Services

The demand for physician services is relevant both as a factor affecting the distribution of physicians and as a measure of the utilization of physician services. Under perfectly competitive conditions, assuming that doctors are of equal ability and wish to maximize income, there would be regional differences in physician-population ratios so long as there are regional differences in demand for services. Under these conditions, however, differences in ratios would correspond exactly to differences in use of services. In equilibrium physicians everywhere would have the same income and provide the same amount of services. No region would suffer from a relative shortage of physician services, if use of services could be taken as an index of need. On the other hand, if there are substantial regional differences in use of services per physician, without corresponding differences in physician income, we may have an indication of maldistribution of physicians.

We shall examine first the factors determining the use of physician services and then the regional use of services in relation to regional physician-population ratios.

The major determinants of the use of physician services are income, age, sex, race, education, place of residence, and possession of health insurance. We need not bother with the use of services as related to sex in a regional analysis. All the other factors tend to be correlated with income. The effect of a person's family income on the number of physician visits per year is an increase of 24 per cent from families with annual incomes under $2,000 to families with incomes of $7,000 and over.[14] In both the "under $2,000" and the "$2,000–$3,999" income groups, the number of visits per person per year is the same, namely 4.6; in the "$4,000–$6,999" income group it is 5.1, and in the "$7,000 and over" group, 5.7. These are averages for entire income groups; the differences are greatest if only children and young persons in different income groups are compared, and smallest in comparisons of only middle-aged persons.

The tendency toward larger number of visits as income rises was detected also in a comparison of cities with different levels of per capita income. The relationship between annual per capita physician visits and per capita effective buying income was studied for eight standard metropolitan areas. The coefficient of correlation was found to be .619 and the coefficient of determination .383. This correlation is significant at the .1 level.

The reasons why persons with high incomes visit physicians more frequently than those with low incomes are no doubt more complex than merely greater ability to pay. The visit differentials should be related to the economic value of good health, perception of illness, and easy access to physicians. These factors are in turn correlated with economic opportunity, education, and place of residence. In 1958–59, whites visited their physicians 35 per cent more often than nonwhites. For families whose head had a college education, the number of visits per person per year was 56 per cent higher than for families whose head had under five years of education and 42 per cent higher than for families whose head had five to eight years of education.[15] For the country as a whole

[14] *Health Statistics*, Series B, No. 19, p. 20.
[15] *Ibid.*, p. 13.

the number of visits per capita in urban areas (towns with 2,500 inhabitants or more) was about 39 per cent higher than in rural farm areas, and about eight per cent higher than in remaining rural areas.

Another factor correlated with income is the possession of health insurance,[16] which tends to increase the demand for medical services. People who possess health insurance tend to have a higher number of physician visits, higher hospital admission rates and higher surgical rates than people without insurance at comparable income levels.[17] This is true not only for services that are covered by insurance but also for those that are not covered. People seem to be more inclined to see a physician if at least some of the potential expense is prepaid.

It has been established that both the physician-population ratio and the number of visits per person increase with income. The question now is whether the demand for services and the relative supply of physicians increase in the same proportion, which would mean an uneven distribution of physicians but an even distribution of work loads. The available data indicate that regional differences in use of physician services are of a lower order of magnitude than differences in physician-population ratios. We computed the number of non-hospital visits per private non-hospital physician by multiplying the regional average number of visits per person by total regional population and dividing by the regional number of physicians. This measure can be taken as a reliable index of work load, for the physicians concerned, unless there are important regional differences in the amount of inhospital visits per physician. The results, presented in Table I, show a large work load differential between the region with the highest and the lowest number of visits. Physicians in the West South Central area (Arkansas, Louisiana, Oklahoma, Texas) have 74 per cent more visits of all kinds outside of hospitals than those in New England and 37 per cent more than those in the Middle Atlantic region. Although New England may be a special case, there is a general tendency for the number of visits to increase with decreases in the physician-population ratio. Apparently, doctors in areas with low physician-population ratios either work longer hours or work more intensively. Contrary to what seemed to be indicated in the earlier comparison of metropolitan centers, the census area comparison indicates a tendency toward a possible redundancy of physicians in high ratio areas and toward a shortage in low ratio areas. However, census divisions are much less homogeneous than metropolitan areas, and

TABLE I

REGIONAL PHYSICIAN WORK LOADS IN THE U. S., 1959

Region	Non-hospital visits per physician	Income per capita	Physicians per 100,000 population	Average income of all nonfederal physicians, 1949
New England	3,817	2,396	113	$ 9,442
Middle Atlantic	4,847	2,543	120	9,574
Pacific	5,284	2,502	112	12,782
East North Central	5,694	2,337	83	12,158
West North Central	5,986	1,978	79	11,961
Mountain	6,042	1,994	84	11,214
East South Central	6,286	1,424	65	11,325
South Atlantic	6,499	1,804	74	11,137
West South Central	6,630	1,764	75	11,794

Sources: *Statistical Abstract of the U.S.*, 1960, Table 6; 1961, Table 419. *Health Manpower Source Book*, Section 10, Table 2; *Health Statistics*, Series C, No. 6, Table 11; *Survey of Current Business*, July 1951, Table 12, p. 19.

[16] Possession of any kind of health insurance is about three times more frequent in families with incomes of $7,000 and over than in families with incomes of under $2,000. See U. S. Department of Health, Education, and Welfare, *Medical Care Financing and Utilization*, Health Economics Series No. 1 (Washington, D. C.: U. S. Government Printing Office, 1962), Table 87, p. 99.

[17] O. W. Anderson and J. J. Feldman, *Family Medical Cost and Voluntary Health Insurance: A Nationwide Survey* (New York: McGraw-Hill, 1960), pp. 73, 183, 194.

TABLE II

PHYSICIAN HOSPITAL WORK LOADS: 1959

Area	Physicians per 100,000 population	Hospital discharges per physician	Patient hospital days per physician
Northeast	149.0	71.3	728
West	128.2	94.7	729
North Central	103.4	113.6	966
South	91.4	127.5	930

Sources: *Health Manpower Source Book*, Sec. 10, pp. 5-6; *Health Statistics*, Series B-No. 32, pp. 18-19.

there is at least the possibility that in higher income areas physicians spent more time taking care of patients in hospitals than in low income areas, which would tend to offset the disparities noted above.

But the opposite is more likely to be the case; that is, in low income regions doctors have probably higher hospital work loads than in high income regions. It was possible to compute the relationships between the number of physicians and the number of hospital discharges and the number of patient hospital days for the four major census divisions of the country. The physicians used for this purpose include all active nonfederal physicians in both private practice and in the hospital service. The results of the computations are presented in Table II, which shows the tendency of higher hospital work loads in areas with low physician-population ratios. The regional unevenness in work loads observed above thus tends to be accentuated when hospital work is taken into account.

We may also conclude that while differences in physical demand for physician services may account in part for the observed regional differences in physician-population ratios, they are not a sufficient explanation of existing disparities in the supply of physicians.

C. Fees for Physician Services

It was shown theoretically that high physician-population ratios in high income areas may be due in part to the fact that fees are based on per capita income rather than on the supply and demand for physician services. It is not necessary, of course, that there be a complete absence of price competition, as was assumed in the theoretical model. A modest amount of price competition may well exist, which would keep fees from rising as much with increases in regional income as they would otherwise.

Some indication that fees are lower in low income areas is implicit in the data presented above in Table I, unless the relationships between regional physician incomes changed considerably between 1949 and 1959. Table I implies that the average income per visit in New England and the Pacific was between 35 and 40 per cent higher than in the southern areas. The East North Central area income per visit was about 20 per cent higher than in the South.

There are no relevant data available on physician fees,[18] but we can remedy this deficiency in part by using another indirect method. The available data on spending and on number of visits make it possible to compute an index of fee differentials between income groups and regions. For the United States as a whole we have for a twelve-month period in 1957-58 an estimated $25 spent on physician services per capita by families earning under $2,000 annually;[19] this income group is reported to have had 4.6 visits per capita during a twelve-month period in 1957-59, which indicates an expenditure of $5.43 per visit. By the same process we obtain an expenditure of $6.93 per visit for the $7,500 and over income

[18] The Bureau of Labor Statistics publishes data on certain types of fees paid by comparable wage and salary earner families in twenty major cities. These data, unfortunately, are not suitable for our purposes, since we are interested in comparisons between income groups.

[19] One third of $75 spent on all health services. See O. W. Anderson, P. Collette, and J. J. Feldman, *Family Expenditure Patterns for Personal Health Service, 1953 and 1958*, Health Information Foundation Research Series No. 14 (New York: Health Information Foundation, 1960), Table 4a, p. 8.

class. The highest income class paid on the average 28 per cent more per visit than the lowest. This is not a measure of price discrimination in the ordinary sense, since it does not refer to a single service but to a composite bundle of services bought by different groups. There may be differences in kind as well as quality between bundles.

There may be more homogeneity in a comparison of expenditures of people living in similar environments. A calculation for farm families only gave the following expenditure differentials: under $2,000 family income, $4.01 per physician visit; $2,000–3,999 income, $4.35 per visit; $4,000–6,999 income, $4.74 per visit; $7,000 and over income, $5.71 per visit.[20] In this case, the highest income group paid 42 per cent more per visit than the lowest. It is noteworthy that the heaviest discrimination occurs between the highest and the next highest income group. A computation by region shows farm families in the South paying $4.13 per visit against $4.82 in the North Central area.

In spite of the necessarily approximate nature of our statistical results, all indications are that prevailing pricing practices contribute to the observed differences in physician-population ratios between areas with different income levels and between urban and rural areas.

V. CONCLUSIONS

It has been shown that high income areas have substantially larger numbers of physicians in relation to population than low income areas. If we assume that in a wealthy country the entire population should have reasonable access to physician services, the implications of this income-related differential deserve careful consideration.

Although there are other factors making high income areas attractive to physicians, there are definite economic incentives working in the same direction. Persons with higher incomes tend to visit physicians more often and have more expensive visits. The larger number of visits per capita in higher income areas tends to justify a certain unevenness in the distribution of physicians in relation to population, but high income areas tend to have more physicians than would be called for by this difference in use of services. The ability to charge higher fees to more well-to-do clients makes possible a high concentration of physicians in high per capita income areas without commensurate loss in physician income. In this sense, the widely defended practice of charging according to ability to pay may have undesirable consequences in terms of the spatial availability of physician services.[21] The availability of services should be considered not only in terms of access to a physician but also in terms of the amount of the physician's time available to the patient on each visit. It was shown that where the physician-population ratio is low, physicians tend to have very large numbers of visits which are almost necessarily brief.

The economic implications of the observed wide regional differentials in physician work loads depend on how health is valued. If we value health from a strictly economic point of view, it is in all probability more advisable to spend a lot of physician time on high income executives, who presumably have a high productivity, rather than on low income unskilled workers. In this case, the observed spatial distribution of physicians may well be optimal, or perhaps there should still be more physicians in high income areas and fewer in low income areas. This approach, however, becomes very awkward when we extend it to the children of various income groups. To the extent that the children have comparable productive potentials, economic

[20] Expenditure data from J. L. Pennock, "Farm Medical Care Expenditures," *Public Health Reports*, April 1958, p. 290; visits are adjusted for rural level.

[21] Within a given area, of course, this practice may make medical care available to some low income groups who would otherwise be deprived of it.

logic suggests that they should have comparable access to medical care.

If we value health from a social point of view and put the same value on the health of all individuals, the economic problem becomes essentially that of adjusting the available supply of physicians as efficiently as possible to the existing need. In the absence of any data on actual need, let us use expressed need, as represented by the number of visits. On these assumptions it can be argued that the existing pattern of physician distribution involves an inefficient allocation of medical manpower resources.

To the extent that wide disparities in regional work loads involve substantial differences in the amount of time a physician can devote to each patient, there is a tendency toward inefficiency. The assumed social values make the social utility of a comparable visit the same for all individuals. From the medical point of view, it is probably reasonable to assume that the effectiveness of a visit does not increase proportionally with the length of the visit beyond a fairly short minimum period. Physician resources made available for long visits are therefore not employed most efficiently. This does not even take into account the fact that where long visits are possible, the visits themselves may be less urgent. Further, where physicians are overloaded, the pressure of time to accommodate many patients may impair a doctor's ability to function with optimal efficiency. Finally, there may be economic inefficiency involved with regard to patients, who may have to waste a great deal of time in the waiting rooms of overloaded doctors. An efficient spatial allocation of physician manpower would therefore be one which tends to equalize the average allowable time per visit between all areas.

A few brief observations are in order also with regard to lack of mobility, which is clearly an important factor in the existing distribution pattern. It is rather doubtful that the existing distribution would be more equal if physicians were more mobile, in the sense that they would be more inclined to maximize either income or some combination of income and leisure. Many of the poorer areas of the country would undoubtedly be even worse off if all physicians made their location decisions with strict economic rationality.

Economic incentives and the lack of mobility also have implications for the improvement of the geographic distribution of physicians. It is unlikely, at least in the short run, that the distribution can be made significantly more even for the country as a whole by simply increasing the manpower output of our medical schools. It is very doubtful whether an increase, perhaps even a substantial increase, in the supply of physicians would automatically solve the distribution problem, as some medical authorities seem to think.[22] Such a measure might result in an increase in the number of physicians in areas where there is a shortage, but this would tend to be accompanied by an even greater increase in areas where there is already a relative excess supply of doctors. This does not promise to be an economical approach to the problem of distribution. On the other hand, there is good reason to believe that the unevenness in the distribution of physicians will diminish automatically, in the long run, along with the tendency toward regional equalization of per capita income.

---

[22] See for instance the statement by Dr. P. R. Hawley in Committee on Medical Care Teaching, *Readings in Medical Care* (Chapel Hill, N. C.: University of North Carolina Press, 1958), p. 27.

Senator NELSON. I would like to submit for the record at this point a letter from E. B. Anderson, Hoffman-La Roche Co.

(The letter referred to follows:)

HOFFMANN-LA ROCHE INC.,
Nutley, N.J., January 2, 1968.

HON. GAYLORD NELSON,
Chairman, Senate Monopoly Subcommittee,
U.S. Senate, Washington, D.C.

DEAR SENATOR NELSON: I want to thank you for the opportunity which you have given us to meet with your staff concerning recent statements which have been issued on the prices paid by certain government authorities for two Roche products. It was most helpful to review together the data which had been accumulated.

Hoffmann-La Roche Inc. shares your belief in the desirability of uniform pricing policies under which government agencies purchasing equal quantities of the same drug product would receive the same price. For some time Roche has had in effect a pricing policy designed to meet this objective, under which any Federal, State or local government which purchases from Roche is charged the same price for the same quantity of the same product.

On December 19, 1967, you issued a public statement and a price table which referred to the prices of several drug products, two of which are sold by Hoffmann-La Roche Inc., chlordiazepoxide hydrochloride and sulfisoxazole. Since several figures contained in the statement and in the survey data released at the same time are incorrect, I am writing to you to provide the correct information, as I am sure you want the record to be as accurate as possible. We also believe that some explanation of other items in the statement and the price table would permit a more meaningful comparison of prices for these two products.

In addition to the statement in the hearing, there was an additional statement by your office released on Decrember 24, 1967, relating to the prices paid by various government purchasers for LIBRIUM capsules (chlordiazepoxide hydrochloride). This later statement, as reported in the *New York Times* of December 25, 1967, and several other papers, also contains some incorrect figures and conclusions which we are sure you would like to see corrected.

In several instances the statements and the table have confused prices for *25 mg.* capsules of chlordiazepoxide hydrochloride with prices for *10 mg.* capsules of the same drug. As a result, the prices for different dosages have been compared, which is obviously misleading.

A. *Chlordiazepoxide hydrochloride*

We have included hereafter the items in the statement and the price table of December 19, and the statement of December 24, relating to chlordiazepoxide hydrochloride which are not in accord with the facts, or which we believe leave an erroneous impression which should be corrected.

1. In regard to the City of Des Moines, the table states that the city pays the highest price of $45.00. While it does not state that Roche sold the product to Des Moines at this price, some persons may have so inferred. Actually, there are no direct sales of chlordiazepoxide hydrochloride to the City of Des Moines by Hoffmann-La Roche Inc. Since Roche has had no direct dealings with Des-Moines for this product, the city's purchase of this product would have been from a wholesaler or retailer. As in the case of Des Moines, the cities of Cleveland and Paterson do not buy directly from Roche. Thus, the prices paid by those cities for Librium are presumably paid to wholesale or retail distributors.

Roche has made some direct sales to county hospitals where these cities are located, but the prices in the table and statement are not the prices which we have charged for these sales. Broadlawns Polk County Hospital (Des Moines) and Highland View Hospital in Cuyahoga County (Cleveland) have purchased Librium *25 mg.* capsules in bottles of 500 for $39.60, which is our price to direct government purchasers which do not buy the 5,000-capsule bottle or in quantities large enough to support a contract. Preakness Hospital in Passaic County (Paterson, New Jersey) has purchased Librium *25 mg.* capsules in bottles of 50 capsules at a price of $4.40 per bottle which is our price for that size bottle. The District of Columbia, we understand, purchases its requirements of Librium from a Veterans Administration depot.

I am sure you understand that the manufacturers of many different types of products cannot sell directly to all end-users, and, as a result, the prices paid by end-users will vary. In addition to direct sales to large-volume users, it is essential that Roche sell through drug wholesalers and retailers in order to serve adequately all of the small-volume purchasers. We, of course, have no control over the prices charged by other persons who distribute our products after we sell to them.

2. The information you received from the City of Philadelphia stated that Philadelphia paid $18.50 for 500 *25 mg.* chlordiazepoxide hydrochloride capsules. Our contract with Philadelphia signed on June 12, 1967, in fact specifies a price of $28.50 per 500 *25 mg.* capsules for a quantity in excess of 100,000 capsules. The price quoted in the statement, namely $18.50, refers to bottles of 500 *10 mg.* capsules for a quantity in excess of 250,000 capsules. In addition, it was stated that New York City paid $28.50 per 500 *25 mg.* capsules of chlordiazepoxide hydrochloride. That was our previous price to New York City. The current contract, which was effective October 1, 1967, calls for $25.50 per 500 *25 mg.* capsules, based on a quantity in excess of 500,000 capsules.

3. The statement noted that the Defense Supply Agency paid $13.20 per 500 *25 mg.* capsules of chlordiazepoxide hydrochloride. Actually, the price of $13.20 per 500 capsules is for *10 mg.* capsules under a contract for over 20 million capsules. Our contract with the Defense Supply Agency for capsules of the *25 mg.* strength calls for a price of $24.00 per 500 capsules in a total purchase of over 3,000,000 capsules.

4. Your information that Philadelphia got a better price than the Veterans Administration for *25 mg.* capsules of chlordiazepoxide hydrochloride is incorrect. As noted above, Philadelphia paid Roche $28.50 per 500 *25 mg.* capsules, based on the purchase of over 100,000 capsules. The Veterans Administration paid $21.75 per 500 *25 mg.* capsules based on a contract purchase of over 5,000,000 capsules. As can be seen, the Veterans Administration pays a lower unit price for capsules of the equal strength because it buys in greater volume.

5. The article also states that "Los Angeles paid a premium price even though it bought three times the quantity Philadelphia bought." This statement is incorrect. As the following table shows, for *10 mg.* capsules Los Angeles paid Roche a lower price based on a quantity purchase of over 1,000,000 capsules than Philadelphia paid based on a quantity of over 250,000; and for *25 mg.* capsules Philadelphia paid Roche $28.50 based on a quantity of 100,000 and Los Angeles paid $25.50 based on a quantity of 500,000:

Philadelphia	
10 milligrams	$18.50
25 milligrams	28.50
Los Angeles	
10 milligrams	16.50
25 milligrams	25.50

6. The prices listed in the article for the cities not covered above are also consistent with Roche's policy of equal prices for government buyers which purchase equal quantities of the same drug from Roche. The cities of Wichita, Kansas; Newark, New Jersey; and Grand Rapids, Michigan, all have paid $39.60 per 500 capsules of the *25 mg.* strength of LIBRIUM. This is the Roche price for a bottle of 500 *25 mg.* capsules, to direct government purchasers who do not buy large sizes, such as the bottle of 5,000 capsules, or annual quantities large enough to justify a contract purchase.

B. *Sulfisoxazole*

The statement and price table of December 19 also contain information on sulfisoxazole which we believe may lead to some erroneous inferences and therefore further explanation.

1. In regard to the City of Des Moines, the table states that the city paid $25.30 per bottle of 1,000 *0.5 gram* sulfisoxazole tablets. Roche does not sell this product directly to Des Moines. Purchases of Roche's sulfisoxazole by the City of Des Moines would have been from a wholesaler or retailer, or Des Moines may have purchased sulfisoxazole from another company.

2. The table quotes a price of $6.90 for a bottle of 1,000 *0.5 gram* sulfisoxazole tablets to the City of Philadelphia. The current Roche contract for sulfisoxazole with Philadelphia, dated June 12, 1967, calls for a price of $9.00 per bottle of 1,000 tablets based on a quantity of 1,000,000 tablets.

3. The prices cited in the table of $8.35 and $8.50 per 1,000 *0.5 gram* sulfisoxazole tablets paid to Roche by the Defense Supply Agency and the Veterans Administration, respectively, are below that paid by Philadelphia because they are supplied under contracts for very much larger quantities.

As you can see in the above items, much of the information upon which you relied regarding the prices for Roche products is inaccurate. We believe that, left uncorrected, these inaccuracies reflect adversely on the reputation and policies of Roche, and upon your Subcommittee. Accordingly, we respectfully request that the records of your Subcommittee be corrected to reflect the facts regarding the Roche prices and policies as stated herein. We appreciate the cooperation which both you and your staff have given us in reviewing this data.

Very truly yours,

E. B. ANDERSON,
Vice President.

Senator NELSON. The committee will adjourn until next Thursday, January 25, in room 457.

(Whereupon, at 1:15 p.m., the hearing in the above-entitled matter was adjourned, to reconvene on Thursday, January 25, 1968, in room 457.)

COMPETITIVE PROBLEMS IN THE DRUG INDUSTRY

THURSDAY, JANUARY 25, 1968

U.S. SENATE,
MONOPOLY SUBCOMMITTEE OF THE
SELECT COMMITTEE ON SMALL BUSINESS,
Washington, D.C.

The subcommittee met, pursuant to recess, at 10:10 a.m., in room 457, Old Senate Office Building, Senator Gaylord P. Nelson (chairman of the subcommittee) presiding.

Present: Senator Nelson.

Also present: Benjamin Gordon, staff economist; James H. Grossman, minority counsel; Susan H. Hewman, research assistant; and William B. Cherkasky, legislative director, staff of Senator Nelson.

Senator NELSON. I now open the hearings of the Monopoly Subcommittee.

Our witness today is Dr. William Comanor.

Dr. Comanor, we appreciate very much your taking the time to come here today to present your testimony to the committee. I have read through your testimony and I think it represents a very thoughtful, valuable contribution to the record of the hearings that we are conducting here on the pricing structure in the drug industry.

I notice that you have a biographical sketch. That will be printed in full in the record preceding your testimony. If you would like to just state extemporaneously your background and your present position, then you may proceed with your statement.

I trust that you have no objection to interruptions for questions?

STATEMENT OF WILLIAM S. COMANOR, PH. D., PROFESSOR, DEPARTMENT OF ECONOMICS, HARVARD UNIVERSITY, CAMBRIDGE, MASS.

Dr. COMANOR. Not at all

(The biographical sketch follows:)

CURRICULUM VITAE

Name: William S. Comanor.
Address:
Residence—10 Fernald Drive, Cambridge, Massachusetts, 02138.
Office—1746 Cambridge Street, Cambridge, Massachusetts, 02138.
Date of birth: May 11, 1937.
Birthplace: Philadelphia, Pennsylvania.
Marital status: Married, no children.
Nationality: United States.

Education:
1951–1955 Central High School, Philadelphia, Pennsylvania.
1955–1957 Williams College, Williamstown, Massachusetts.
1957–1959 Haverford College, Haverford, Pa., A.B. in Economics.
1959–1963 Harvard University, Cambridge, Mass., Ph. D. in Economics.
1963–1964 London School of Economics, London, England.
Thesis title: "The Economics of Research and Development in the Pharmaceutical Industry."
Fields of interest:
Industrial Organization and Public Policy.
Economics of Science and Research.
Statistics and Econometrics.
Economic Theory.
Honors and Fellowships:
Sophomore Honors, 1957.
High Honors in Economics, 1959.
Phi Beta Kappa, 1959.
Harvard University Fellowship, 1959.
National Science Foundation Fellowship, 1963.
Sloan Foundation Research Grant, 1966.
Professional experience:
Teaching Fellow, Harvard University, 1961–1963.
Resident Tutor, Quincy House, Harvard University, 1961–1963, 1964–1965.
Economist, Caterpillar Tractor Co., Peoria, Ill., Summer 1960.
Economist, United Research, Inc., Cambridge, Mass., Summer 1961.
Research Assistant, Professor Richard E. Caves, Harvard University, Summer 1963.
Instructor in Economics, Harvard University, 1964–1965.
Special Economic Assistant to the Assistant Attorney General, Antitrust Division, U.S. Department of Justice, Washington, D.C., 1965–1966.
Assistant Professor of Economics, Harvard University, 1966.
Consultant, Booz, Allen, and Hamilton, Inc., 1966–1967.
Consultant, Antitrust Division, U.S. Department of Justice, 1966.
Publications:
"Research and Competitive Product Differentiation in the Pharmaceutical Industry in the United States," *Economica*, November 1964.
"Review and Technical Change in the Pharmaceutical Industry," *Review of Economics and Statistics*, May 1965.
"The Drug Industry and Medical Research: The Economics of the Kefauver Committee Investigation," *Journal of Business*, January 1966.
"Competition and the Performance of the Midwestern Coal Industry," *Journal of Industrial Economics*, July 1966.
"Vertical Mergers, Market Power, and the Antitrust Laws," *American Economic Review*, May 1967.
"Advertising Market Structure and Performance," *Review of Economics and Statistics*, November 1967 (with Thomas A. Wilson).
"Market Structure, Product Differentiation, and Industrial Research," *Quarterly Journal of Economics*, November 1967.

Senator NELSON. Go ahead, Doctor.

Dr. COMANOR. My name is William S. Comanor. I am presently assistant professor for economics at Harvard University. I received my A.B. from Haverford College in 1959 and my Ph. D. from Harvard in 1964. During the academic year 1965–66, I served with the Department of Justice as Special Economic Assistant to the Assistant Attorney General in charge of the Antitrust Division. I currently serve as a consultant to the Antitrust Division of the Department of Justice.

My field of special interest within economics is industrial organization, problems of competition and monopoly, and government policy toward business. For sometime now, I have been especially concerned with the economics of the pharmaceutical industry. My Ph. D. thesis was entitled "The Economics of Research and Development in the Pharmaceutical Industry." From my research, I have published three

articles dealing with the economics of the drug industry. The first is entitled "Research and Competitive Product Differentiation in the Pharmaceutical Industry in the United States." [1] It appeared in the 1964 issue of Economica. The second is entitled "Research and Technical Change in the Pharmaceutical Industry," [2] and appeared in the May 1965, issue of the Review of Economics and Statistics. The third is entitled "The Drug Industry and Medical Research: The Economics of the Kefauver Committee Investigation," [3] and appeared in the January 1966, issue of the Journal of Business, published by the University of Chicago.

I have copies of all of these articles. You may wish to include them in the record.

Senator NELSON. The articles you have mentioned will be printed in the record at the conclusion of your testimony, along with the supplemental study you have furnished the committee, entitled "Advertising Market Structure and Performance," [4] which I understand is to be published.

Dr. COMANOR. That is right. I will mention that.

Senator NELSON. That study will also be printed at the conclusion of your testimony.

Dr. COMANOR. I would like to report this morning on some findings from a study which Prof. Thomas A. Wilson, of the University of Toronto, and I have been carrying out on the relationship between advertising and profit rates in the American economy, and then I want to comment on some of the implications of the study for competition and public policy toward the pharmaceutical industry.

The statistical results are contained in a paper entitled "Advertising Market Structure and Performance," which will be published in the November, 1967, issue of the Review of Economics and Statistics. I am afraid that although economists advocate efficiency for others, they are frequently as lax in their own operations as anyone else, and the November, 1967, issue of this review is not scheduled for publication until February and March of the current year.

THE STATISTICAL FINDINGS

The core of this paper is a statistical analysis of the relationship between advertising and profit rates across a wide range of consumer goods industries. Our analysis was limited to this sector of the economy for it is only here that advertising seems to have a major competitive impact.

Even among consumer goods industries, advertising seems far more important in some industries than others. In most industries, advertising expenditures account for a relatively low proportion of total revenues while in a handful of others, they reach very high proportions. Thus, in the 41 consumer goods industries examined in our study, 25 had advertising-sales ratios below 3 percent, eight had ratios between 3 percent and 6 percent, and only six industries had ratios which exceeded 6 percent.

[1] See article, p. 2069, infra.
[2] See article, p. 2078, infra.
[3] See article, p. 2086, infra.
[4] See article, p. 2092, infra.

In the latter group, perfumes had an advertising-sales ratio of 15 percent, cereals and drugs were approximately 10 percent each, soap 9 percent, beer 7 percent, and soft drinks slightly more than 6 percent.

At lower levels, cigarettes and wines had advertising-sales ratios of about 5 percent.

These findings suggest that the economic effects of advertising expenditures are not likely to be felt generally throughout the economy but rather are limited to a small subset of industries.

Senator NELSON. Doctor, I notice that you comment on this later on in your testimony, but the figure that you use for expenditures for advertising of drugs is 10 percent of sales.

What you are saying is that the total expenditure for advertising for drugs equals 10 percent of the total sales of drugs, is that correct?

Dr. COMANOR. That is right.

Senator NELSON. And in this figure, you are using both prescription and nonprescription; that is, proprietary drugs.

Dr. COMANOR. That is right.

Senator NELSON. What is the total amount spent on prescription and proprietary drugs?

Dr. COMANOR. I have not looked at the most recent figures, it is something in the order of $2.5 billion, $3 billion, but this could be checked.

Senator NELSON. This figure confuses me. My memory is that we have had testimony that somewhere around $4 to $4.5 billion is spent on prescription drugs alone in this country.

Dr. COMANOR. It has been some time since I looked at the specific figure that you are asking. I may be completely wrong. I would like to check the numbers before I answer the question.

Senator NELSON. Do you have those figures in your study?

Dr. COMANOR. The figures which you ask are not contained in the printed paper which I am presenting to you, but I have them in Cambridge. I can certainly send them to you.

Senator NELSON. I would appreciate your getting these figures for the committee, because it seems to me quite a critical statistic. If I recall the testimony correctly, $4.5 billion is now spent on prescription drugs; that is one figure. I would like to know what the figure is on proprietary drugs. Then the next aspect of the question is, it seems to me, the critical aspect of advertising in the drug industry; that is, advertising for prescription drugs, which is directed only to the medical profession. I would like to know what percentage the advertising cost as a percentage of sales for prescription drugs. Again the figure that we have here is that $800 million is spent on advertising of prescription drugs, not counting detail men; $800 million as a percent of $4.5 billion is a much higher advertising figure than the figure you use of 10 percent.

I may be wrong in my recollection on the statistics.

Dr. COMANOR. The 10-percent figure applies both to proprietary as well as prescription drugs. As I note later in my statement, it applies to the average values for the years 1954 through 1957. So we are dealing with a period over 10 years ago.

Senator NELSON. Your statistics apply to what period?

Dr. COMANOR. They apply to average values for the years 1954 through 1957. I state later in my testimony precisely why this period was chosen. They are not the most recent figures that one could find on this particular percentage in the pharmaceutical industry.

Senator NELSON. The reason it seems to me that it would be valuable to have the statistics separated out with a separate set of statistics for proprietary and prescription drugs is that they present two very different problems in terms of the difficulty in entering the marketplace that you talk about. It seems to me that it is one kind of problem to get into the marketplace and compete in the sale of a prescription drug when all of the advertising of that drug is aimed solely at the medical profession and quite another problem to get into the marketplace with a proprietary drug where all the advertising, or almost all of it, is aimed at the general public.

Dr. COMANOR. I agree with you completely. I think these two separate industries create very different problems for public policy. This figure was used because the data was obtained from the Statistical Division of the Internal Revenue Service and the IRS publishes data for the two industries combined. The only tax return data that are available to the general public for this particular industry combine both proprietary and prescription drugs. So that although for public policy purposes, I agree with you completely, for the purposes of our study, we were forced to use the combined industry.

Senator NELSON. That is because the profit figures——

Dr. COMANOR. The profit figures also apply to the combined industry. This is the way the IRS publishes tax return data.

Senator NELSON. But the statistics are available, are they not, for the total amount of proprietary drugs sold and the total amount of prescription drugs sold?

Dr. COMANOR. That is right, but we wanted to undertake a study dealing with consumer goods industries generally throughout the economy.

Senator NELSON. I realize your study was not directed at precisely all of the points that are being raised here. But for purposes of the record, if the statistics are easily available, I would appreciate it if you would submit the most current statistics that you can get on total sales of prescription drugs and what percent of sales is represented by advertising for prescription drugs and the same for proprietary drugs. I think it will probably show that the percentage of advertising to sales of prescription drugs is much higher than this average of 10 percent that you have given us for all drugs.

Dr. COMANOR. I will try to get the statistics for you.

Senator NELSON. If it is easily available to you, if you would submit that for the record in the next week or 10 days so we could print it at this stage of the testimony, I would like that.

Dr. COMANOR. Fine.

(The material referred to, subsequently received, follows:)

HARVARD UNIVERSITY,
DEPARTMENT OF ECONOMICS,
Cambridge, Mass., February 8, 1968.

Senator GAYLORD NELSON,
Chairman, Monopoly Subcommittee,
Select Committee on Small Business,
U.S. Senate, Washington, D.C.

DEAR SENATOR NELSON: During my testimony before the Monopoly Subcommittee on January 25, 1968, you requested that I submit some additional information to be included in the record.

In the 1963 Census of Manufacturers, Pharmaceutical Preparations (Standard Industrial Classification 2834) were divided into two classes: ethical and proprietary. Domestic shipments of ethical preparations were $2,054,897,000, while domestic shipments of proprietary preparations were $836,166,000. I am afraid, however, that total advertising expenditures in these two industries are not available from public sources. For the same year, Internal Revenue Service statistics report that total business receipts in the drug industry (IRS Minor Industry 2830) were $4,505,870,000 while total advertising expenditures for the same firms were $405,007,000, or 9.0% of sales. For the following year, 1964, the IRS listed a smaller number of firms in this industry, and business receipts for these firms were $4,175,331,000, while total advertising outlays amounted to $461,648,000, or 11.1% of sales.

The drug industry, as defined in IRS statistics, include sales of both ethical and proprietary products. Furthermore, the IRS statistics include advertising outlays and revenues for entire firms even though some firms may have considerable operations which are not related to the productions and distribution of drugs. Census statistics, on the other hand, are based on data gathered at the establishment level, and therefore include only business operations which deal specifically with a particular industry.

As I noted in my testimony, the Statistics Division of the Internal Revenue Service accepts the definition of advertising which is made by individual firms in their tax returns. Different firms, moreover, appear to list different categories of expense under different headings, so that gifts for doctors and medical students are included in advertising expense by some firms while others list these outlays under promotional expense. Furthermore, I understand that contrary to my statement on January 25, the cost of samples are sometimes included in advertising expense. Thus, it appears that reported outlays on advertising understate the true volume of advertising and promotional effort in the drug industry even apart from the salaries paid to detail men.

Sincerely,

WILLIAM COMANOR.

Dr. COMANOR. In most industries, even in the consumer goods sector, advertising probably plays a relatively minor role, although it appears to play a major role in a few others.

One further feature of these statistics is that the subset of industries with high advertising-sales ratios is comprised entirely by those which produce consumer nondurables rather than consumer durables. Industries which produce electric appliances or radio and television sets spend relatively little on advertising as a proportion of total sales. And yet, it is precisely these industries which are generally considered to produce highly complex products about which consumers might need considerable information.

In the statistical analysis, we examine the joint effect of advertising, together with a number of other market structure variables, on profit rates. The additional variables are the degree of market concentration, the rate of growth of demand, and estimates of two entry restricting factors: the extent to which production economies of scale exist in the industry, and the total amount of capital required for entry. These additional variables were introduced into the analysis although our

primary attention was directed at the impact of advertising outlays because they are generally considered to influence profit rates and we wished to examine the net effect of advertising on profits after the influence of these other factors had been accounted for.

The statistical analysis is founded largely on tax return data which are compiled by the statistical division of the Internal Revenue Service. It is based on industries as defined by the IRS and refers to average values for the years 1954 through 1957. Averages were taken so that the results would not be influenced by the particular business conditions which happened to exist in a given year, and the years chosen refer to a complete business cycle.

The statistics on profit rates and advertising-sales ratios which are used in the empirical analysis are presented in table A–1 attached to my statement.

(The statistics referred to follow:)

TABLE A–1.—AVERAGE PROFIT RATES AND ADVERTISING-SALES RATIOS IN 41 CONSUMER GOODS INDUSTRIES, 1954–57

[In percent]

	Profit rate	Advertising sales ratio
1. Soft drink	10.0	6.2
2. Malt liquors	7.2	6.8
3. Wines	7.3	5.2
4. Distilled liquors	5.0	2.1
5. Meat	4.6	.6
6. Dairy	7.9	2.2
7. Canning	6.4	2.9
8. Grain mill products	7.0	1.9
9. Cereals	14.8	10.3
10. Bakery products	9.3	2.9
11. Sugar	5.8	.2
12. Confectionery	10.6	3.5
13. Cigars	5.3	2.6
14. Cigarettes	11.5	4.8
15. Knit goods	3.8	1.3
16. Carpets	4.5	2.0
17. Hats	1.6	2.2
18. Men's clothing	5.9	1.2
19. Women's clothing	6.1	1.8
20. Millinery	−1.3	.8
21. Furs	5.7	1.0
22. Furniture	9.7	1.5
23. Screens and venetian blinds	9.3	1.6
24. Periodicals	11.7	.2
25. Books	10.1	2.4
26. Drugs	14.0	9.9
27. Soaps	11.7	9.2
28. Paints	9.9	1.5
29. Perfumes	13.5	15.3
30. Tires and tubes	10.2	1.4
31. Footwear	7.6	1.5
32. Handtools	11.4	4.2
33. Household and service machinery (not electrical)	7.3	1.9
34. Electrical appliances	10.3	3.5
35. Radio, TV, and phonograph	8.8	2.2
36. Motorcycles and bicyles	5.2	1.1
37. Motor vehicles	15.5	.6
38. Instruments	12.0	2.0
39. Clocks and watches	1.9	5.6
40. Jewelry (precious metal)	5.3	3.2
41. Costume jewelry	1.4	4.0

Dr. COMANOR. A significant correlation exists between these two variables. While high profit rates may be and frequently are associated with relatively low advertising-sales ratios, such as instruments which include cameras, industries with high advertising-sales ratios

almost without exception have relatively high profit rates. And this refers specifically to the six industries that have advertising-sales ratios which exceed 6 percent.

Thus, the average profit rate for those six industries was 11.9 percent which is 65 percent greater than the average return for the 35 remaining industries of 7.2 percent.

Senator NELSON. When you say "average profit rates," are you talking about after-tax profits?

Dr. COMANOR. I am talking about after-tax profits.

While motor vehicles has the highest profit rate of any industry in the sample, and at the same time has a relatively low advertising-sales ratio, the next three most profitable industries—cereals, drugs, and perfumes—are all those with very high levels of advertising expenditures.

Senator NELSON. Does the automobile industry fit into the picture you describe later of an industry that has some very difficult entrance barriers—that is, the vast investment it takes to get in?

Is advertising a factor there?

Dr. COMANOR. I think most economists would agree that the automobile industry has very high entry barriers, but that these are not accounted for by advertising expenditures as much as by the important scale economies which exist in the industry, as well as by the very high absolute capital requirements. In addition, the current franchise system does tend to restrict entry. While important entry barriers exist, I think most economists would not stress the effect of advertising in this particular industry, although firms in the industry spend very large absolute sums on advertising, since after all, it is one of our largest industries. But these large absolute sums account for a relatively small percentage of sales.

Senator NELSON. Would you agree that there is a distinction between the automobile and the prescription drug industry on at least a couple of grounds relative to the advertising question, the automobile industry having a relatively low advertising ratio to sales, with the drug industry having a high advertising ratio to sales? Would you agree that one of the reasons for the distinction there is that there is a multiplicity of drugs and the person who makes the determination about what drugs will be used and what brand will be used is the doctor, whereas in the automobile industry, there are only four major companies in this country and the consumer, by looking at the product, makes the decision about what he will buy?

Dr. COMANOR. While this factor is certainly important, I would emphasize a somewhat different factor, which is that in the motor vehicle industry, factors such as high absolute capital requirements inhibit entry so that the auto firms do not feel compelled to allocate further sums to restrict the entry of new firms.

In addition, while advertising promotes product differentiation in the automobile industry, there are various other factors besides advertising which are at least as important, if not more important, in creating product differentiation, such as the annual model changes.

In the pharmaceutical industry, where products which embody the same chemical compound but which are produced by different firms are now similar to one another, advertising plays a much more im-

portant role. It is in this context that it is interesting to note that high advertising outlays generally are concentrated in the nondurables field, where product differences are probably much less. Advertising is very high, for example, in the detergent industry, although detergents are probably less different from one another than automobiles.

In addition, we know that in the proprietary drug field, there are very high advertising outlays for products such as aspirin, although aspirin is aspirin.

Therefore, what we find is that advertising expenditures tend to be highest where real product differences are not pronounced. And advertising tends to be relatively low where real product differences in fact exist.

One might conclude that if there are real product differences, advertising does not play the role that it does play when these differences are absent. It is in this way that I think we can interpret the statistics which indicate that advertising-sales ratios are higher in the consumer nondurable sector than in the consumer-durable sector.

Senator NELSON. I realize there are a lot of factors involved in this, but there are several differences. There is one fundamental difference between the automobile and the drug industry; that is, that the consumer in the one case, automobiles, is qualified to make a judgment of his own to decide whether he likes the style of the car, what horsepower he wants, how many miles per gallon it will travel, what size automobile he wants, and then based upon his experience in buying cars, he is qualified to make his own judgment. However, he has no qualifications for making any judgment whatsoever about a prescription drug.

So they are fundamentally different problems, are they not?

Dr. COMANOR. I agree, they are quite fundamentally different problems. At the same time, we can and do find high advertising sales ratios in industries where, in fact, the consumer does make the purchasing decision—such as cereals and perfumes and soaps. So I would not stress this particular factor as the primary reason why advertising outlays are high in this industry, although I think that the fact that the doctor is the purchasing agent for the final consumer plays a very crucial role on the type of advertising which is undertaken and the effectiveness of these advertising outlays.

Senator NELSON. I was not suggesting that this factor that I mentioned was the most important, or one of the most important, but that it is a factor which differentiates the two products. Please proceed.

Dr. COMANOR. These preliminary observations were corroborated by our statistical analysis where the influence of the additional market structure variables was also considered. We found that advertising has a statistically significant impact on industry profit rates and that this effect is stronger than that of any of the other variables examined.

Furthermore, the magnitude of the effect is surprisingly high. Industries with high advertising outlays earned, on average, profit rates which exceeded those in other industries, after correcting for the other variables, by nearly 4 percentage points. This difference represents, moreover, a 50-percent increase in profit rates. Since profits represent the difference between prices and costs, and since advertising outlays is a cost to the firm making the outlay, these findings suggest

that high advertising expenditures have a double effect on price levels. Not only do they represent higher costs, but also they are associated with higher price-cost margins.

ADVERTISING AS A BARRIER TO ENTRY

The empirical evidence described above indicates that industries which spend substantial proportions of their revenues on advertising tend at the same time to earn high rates of return, and I want now to examine some economic factors which might explain this result.

In a free market economy, high profit rates, which are not accounted for by differences in rates of growth of demand or by the degree of risk, will not persist for a prolonged period of time in the absence of factors which restrict the entry of new firms. If high rates of return in an industry, in firms will enter the market, and this will generally have the effect of driving prices down to more competitive levels. High prices and profits suggest, thereby that specific factors are present which restrict the entry of new firms.

Senator NELSON. The economists who have testified in behalf of the Pharmaceutical Manufacturers Association have argued that the high profits of the drug industry were the result of high risks.

Do you have any observations to make about that?

Dr. COMANOR. Yes; I do. I have read parts of the testimony presented by these economists .

It seems to me that to consider this, one would have to ask what we mean by risk and in how we can measure it. Risk involves some notion of unexpected changes in the value of the firm or unexpected changes in profitability. But this cannot be measured simply by the variability in profit rates across firms in an industry. After all, there are generally differences among profit rates. They may measure accustomed and expected and traditional differences among firms. They may measure systematic as opposed to unexpected differences. Different firms in an industry may have different profit rates for a number of reasons: because of economies of scale, because of other differences in efficiency. because they operate and sell in different segments of the market, or because of differences in consumer acceptance or product differentiation.

It seems therefore that the variability of profit rates across firms in an industry cannot be used to measure risk because these changes may be accustomed and traditional rather than unexpected and the essence of risk is the inability to predict what will happen in the future.

In addition, risk, however measured, is likely to be very different for firms of different sizes, and the simple variability across firms does not account for these differences.

This is another reason why simple variability across firms is not an adequate measure of the degree of risk in an industry.

It seems to me that a much better measure might be the intertemporal difference and the variability of profit rates, across time, especially after correcting for size of firm. I note from testimony given recently that when risk is measured in this manner, the degree of risk for the drug industry tends to be very low.

These statistics were presented to Dr. Mueller's testimony. Therefore, I think it is difficult to argue, and I certainly would not, that the drug industry is a high risk.

Senator NELSON. Did you read Dr. Mueller's testimony?

Dr. COMANOR. Yes, I did.

Senator NELSON. Do you agree with him on the point he makes on this question?

Dr. COMANOR. Yes, I do.

In the consumer goods industries where advertising is relatively important, we frequently observe unbranded products which sell at prices substantially below those of highly advertised products even though there may be little real difference between them. Because products which are little advertised must sell at far lower prices than those of their established rivals, the latter can raise prices above costs and earn high profits without fear of enticing the entry of new firms and of the resulting effect on price levels.

High advertising outlays create effective entry barriers through a number of routes. In the first place, high current levels of advertising expenditures create additional costs for new entrants which will generally exceed those for established firms. Because of buyer inertia and brand loyalty, more advertising messages per prospective customer must be supplied to induce brand switching as compared with repeat buying. And therefore, new entrants must be prepared to supply more advertising messages per prospective customer than do their established rivals.

A further disadvantage faced by new entrants is that they must spend nearly as much in total advertising and other forms of promotion as existing firms if their products are to compete with established and well-known products. Since new entrants are generally small, this total expense must be spread over fewer units of output, and therefore the cost per unit may be quite high. Larger established firms, on the other hand, have the advantage of being able to spread this cost further and often spend considerably less per unit sold. Thus, where advertising is important, new entrants are frequently found to bear higher unit advertising costs as compared with their larger rivals.[5] And this factor again enables established firms to set relatively high prices without fear of the effective competition of new entrants.

Furthermore, there is evidence that the cost per advertising message declines as the number of messages purchased increases.[6] New entrants, thus, face higher unit costs which contribute further to their relative disadvantage.

Finally, the need to spend considerable funds on advertising will raise the amount of capital required for entry into the market. Since new firms frequently find it difficult indeed to raise large amounts of

---

[5] In the automobile industry, for example, the two smaller firms during the 1950's were found to spend more than twice as much on advertising per car sold as did either Ford or General Motors. Between 1954 and 1957 Studebaker and American Motors spent annually on national advertising approximately $64.04 and $57.89 per automobile sold while General Motors spent $26.56 per unit and Ford spent $27.22 per unit. Chrysler assumed an intermediate position, spending $47.76 per unit. Leonard W. Weiss, "Economics and American Industry," p. 342.

[6] The extent of discounts given to large advertisers is documented in *Federal Trade Commission* v. *The Procter & Gamble Co.*, brief for the Federal Trade Commission in the Supreme Court of the United States, December 1966, pp. 12–13.

capital, high advertising expenditures by existing firms, and thereby high advertising requirements for new entrants, will further restrict the rate of entry.

For all of these reasons, heavy advertising expenditures serve to create substantial entry barriers. They act as an important restriction on competition and permit established firms to charge high prices and earn high profit rates without fear of the competitive consequences. And the statistical findings presented earlier merely describe the extent of the effect of advertising on competition.

THE PROBLEM OF CASUALITY

Throughout the analysis, we have assumed that the direction of casuality runs from advertising to high profit rates, and it is necessary to consider whether the reverse could be the case. This is especially so since a plausible argument could be made that, since advertising reflects the discretionary behavior of firms, high profits could lead to high advertising. In other words, high profits could be "spent" on advertising.

There are a number of factors, however, which suggest that the appropriate direction of casuality is in fact from advertising to profits. The first is that if profits were "spent" on advertising, which is after all an expense to the firm, higher measured profit rates would be associated with lower advertising expenditures, and lower measured profit rates, with higher outlays. But we have observed a positive rather than a negative relationship between these two variables.

A second and even more important reason stems from the fact that the amount of advertising expenditures depends on many factors besides the whims of individual managers.

Product and market characteristics make advertising a more profitable activity in some industries than in others, and there are few reasons for believing that managers in some industries are better equipped to take advantage of their opportunities than those in others.

In fact, if we assume that managers make their decisions on advertising burgets so as to maximize profits, then the differences between industries which we observe reflect not the discretionary behavior of individuals but rather the particular product and market characteristics of member firms in the industry.

The figures presented above, which deal with the average values across a 4-year period, may be interpreted in fact as describing the optimal levels of advertising expenditures in the particular industry. And there is no reason to believe that firms with higher profit rates will have higher optimum advertising expenditures than firms with lower profit rates but similar market conditions.

This latter consideration is an important one for interpreting the statistical results presented earlier. Since the analysis focuses on interindustry differences in advertising outlays, and since these differences are more likely to reflect the product and market characteristics in the industry rather than the peculiarity of individual managers, these results have few behavioral connotations. They do not imply that an industry can earn higher profit rates simply by spending more on

advertising. They do not suggest that firms which currently spend little on advertising should double their advertising budgets if their objective is to obtain higher returns. But rather, they describe the fact that firms and industries with higher optimum advertising expenditures will and do earn higher rates of return than firms in less advantageous market positions.

THE POSITION OF THE DRUG INDUSTRY

Although this study refers to the entire consumer goods sector of the economy, we can note that the drug industry is one of those in the sample with both high profit rates and high advertising-sales ratios. During the period between 1954 and 1957, this industry stood in third place out of 41 in both respects. As defined in this analysis, however, the drug industry refers to both prescription and proprietary products.

In the prescription drug industry, advertising accounts for only a portion of the firm's total selling and promotional budget. Detail men and other forms of direct selling also are important, and therefore figures on advertising alone understate the total effort in persuasion which is carried on by the drug companies.

Senator NELSON. In your figures on advertising, you did not include any of the costs of the detail men?

Dr. COMANOR. That is correct.

Senator NELSON. What did you include?

Dr. COMANOR. We included expenditures on advertising as reported in tax returns.

Senator NELSON. You reported what the industry claimed as advertising and what the IRS accepted as advertising?

Dr. COMANOR. That is correct.

Senator NELSON. Do you know what that included?

Dr. COMANOR. That includes expenditures on printed advertisements, direct mail advertisements, and also advertisements in medical journals. It does not include salaries for detail men, for samples, or things of this sort.

Senator NELSON. It did not include samples?

Dr. COMANOR. It did not include samples.

Senator NELSON. Even though samples had some advertising material with it?

Dr. COMANOR. I think that is correct.

Senator NELSON. Did you include other expenditures such as material furnished to medical students, stethoscopes, that sort of thing?

Dr. COMANOR. I am not absolutely certain, but I think not. I would have to check with the IRS before I could answer these questions completely.

Senator NELSON. I do not know what these expenditures amount to, but we have had testimony that medical students get a considerable amount of material, some advertising and promotional, some of them stethoscopes and so forth.

Do the firms deduct this as a business expense?

Dr. COMANOR. This is surely a business expense. I am not certain whether it is included in these advertising figures. My wife is a medical

student, and I am forced to admit that she has received a stethoscope and a black bag.

Something that you might be interested to know is that one of the drug companies provides every second-year medical student with a black bag which is plastic. The same company in their fourth year provides the medical student with a black bag of real leather. She is now going around with her plastic one.

Senator NELSON. They have somewhat less confidence in the prospects for the second year student than the senior student?

Dr. COMANOR. That must be it.

Senator NELSON. Would it be easy for you to check and see whether these items include this advertising expense in the tax column?

Dr. COMANOR. One could check with the IRS on this question.

Senator NELSON. If you could, along with the other material you are going to submit, it would be useful to put it in the record.

Dr. COMANOR. I will do this.[7]

Advertising generally is particularly effective in situations where consumer uncertainty or ignorance about the relative merits of competing products is relatively high. It is not so much that physicians or other consumers consciously accept the blandishments of an advertising message as much as they consider the purchase or prescribing of an advertised product as the "safe" course of action.

When considerable uncertainty exists, the choice of a highly advertised product often represents minimizing the risk to the consumer that the product will do the job for which it is being purchased. And what becomes important is not the content of the advertising message but simply the fact that the product is well advertised. Thus, products are frequently described as "advertised in Life." [8]

The effect of advertising on entry barriers in the pharmaceutical industry is strengthened by an important interaction which exists between research and selling expenditures. The relatively high level of research in this industry has led to a rapid rate of new product introduction throughout the postwar period. During the decade ending in 1960, over 3800 new products or dosage forms were introduced into U.S. pharmaceutical markets. Of this total, slightly more than 11 percent were new single chemical entities.

Given the rapid pace of new product introduction, it is not surprising that doctors have been open to the persuasion and influence of the drug companies. At a 1959 medical conference on the evaluation of new drugs, it was reported that "physicians are frightened, confused and puzzled by advertising material which pushes as many as a thousand new drugs or combinations of drugs every year * * * several practitioners at the session said they felt a void of information about the proper use of new drugs." [9] The simple pace of new product introduction has been a major factor which has increased the effectiveness of heavy advertising and promotional expenditures.

Furthermore, most new products have received patent protection. Thus, in 1961, of the 656 single chemical entities which were used for

[7] See letter, p. 2046, supra.

[8] This view is similar to that taken in Donald F. Turner, "Advertising and Competition," speech delivered before the Briefing Conference on Federal Controls of Advertising and Promotion, June 2, 1966, pp. 3–4.

[9] Herman Somers and Anne Somers, "Doctors, Patients, and Health Insurance," p. 100.

therapeutic purposes, 377 were sold only by a single firm, which amounted to about 57 percent of the total.[10] The impact of the patent system, however, has not been to create tight monopoly positions, since patented products are often highly substitutable and compete with one another. But rather, it has been to foreclose, to a great extent, rivalry between identical chemical entities or standardized commodities about which price competition might develop. It has strengthened and encouraged a form of chemical product differentiation. In the majority of cases, therefore, advertising has been able to exploit and emphasize the chemical differences which do in fact exist among products.

When doctors are forced to choose between different chemical compounds which purportedly do the same thing, a large measure of uncertainty is certain to exist. Even though a doctor might believe that different compounds have similar therapeutic effects, he is never quite sure, and the prudent course of action is to prescribe the drug which has become well-known. And becoming well-known is, of course, the function among other things of the level of advertising and promotional expenditures.

Mr. GORDON. Did you say the purpose of the advertising and promotional activities is to prevent price competition?

Dr. COMANOR. After the experience of the drug industry with penicillin, which was introduced in the early postwar period, when the price dropped precipitiously in the course of 5 or 8 years—as I recall, the price of penicillin for a standard dosage dropped from $20 in 1943 to on the order of five cents by 1950——

Senator NELSON. $20 for——

Dr. COMANOR. I think it was 100,000 units.

Senator NELSON. From $20 to what?

Dr. COMANOR. The price dropped from $20 for 100,000 units in 1943 to 4.5 cents in 1950. The price of streptomycin behaved similarly.

At the prices which finally resulted, profits were clearly quite low.

Senator NELSON. What accounted for the drop?

Dr. COMANOR. Penicillin is a standardized commodity. A large number of new firms entered into the market, vigorous competition took place, and prices were pushed down to costs.

Senator NELSON. Was a profit made on it?

You say the prices were pushed down to cost. That would mean no profit.

Dr. COMANOR. It is unclear as to just what the profit margin was for penicillin. It may very well have been that profits were low and even nil. No figures are available on this question.

This experience, I think, had a sharp impact on the industry. It became clear that profits in the future would depend on introducing specialized products, on restricting the growth of standardized commodities, and thereby restricting effective price competition. Price competition seems to demand a number of firms who sell the same chemical compound, for it is only in this setting that the purchaser may focus on price rather than the therapeutic effects of a product. It became clear that some means of differentiating the product was re-

[10] Hugh P. Walker, "Price Levels and Market Power in the Ethical Drug Industry," paper presented at the December 1967 meetings of the Econometric Society, Washington, D.C., table 3.

quired. We note that throughout the decade of the 1950's, large numbers of new products were introduced. Most of these were protected by patents. And also, advertising expenditures grew rapidly. We can understand the incentive for both research and advertising in the drug industry in terms of the aim to create and promote product differentiations which would prevent the development of standardized commodities and which would prevent the type of price experience which occurred in penicillin and streptomycin.

Senator NELSON. Thank you.

Mr. GORDON. The product differentiation you just talked about is the chief barrier to entry into the various therapeutic categories, is that correct?

Dr. COMANOR. That is correct. It is the role of both advertising and research in this industry to promote a form of product differentiation which then serves to restrict entry and to restrict the development of price competition.

In this manner, the patent system interacts with the high level of advertising and promotion by insuring that the competition which exists in this industry will focus on chemically differentiated products, and this factor increases substantially the effectiveness of the selling efforts in this industry.

Were the patent system not to exist, the effectiveness of these efforts is likely to be reduced. If new entrants were free to produce the same chemical compound as established firms, the effectiveness of advertising in constructing barriers to entry would be lessened, and we might expect to find more stringent constraints placed upon the price behavior of the leading firms in the industry.[11]

Even in the absence of chemical differences among products, however, considerable uncertainty remains as to whether different brands of the same compound are equally good. It is claimed that tradename products which are sold by the larger companies are manufactured under quality controls which produce higher quality products. While there is no readily available evidence on quality control expenditures for pharmaceutical firms in the United States, statistics are available for Canada. Twenty-one large firms, many of which are U.S. subsidiaries, reported to the Canadian Government that quality control expenditures accounted for only 1.2 percent of total sales or approximately 3.6 percent of the cost of goods sold.

Senator NELSON. What is the distinction? 1.2 percent——

Dr. COMANOR. Of total sales.

Senator NELSON. I understand the 1.2 percent of total sales and approximately 3.6 percent of the cost of goods sold—what precisely do you mean?

Dr. COMANOR. Production costs. You can translate cost of goods sold as production costs, of the total production costs of the product, quality control accounts for only 3.6 percent. After these costs are advertising, promotion, profits, to make up total revenues.

Senator NELSON. So when you are talking about cost of goods sold, you mean the total cost of producing the product?

---

[11] Restrictive Trade Practices Commission, "Report Concerning the Manufacture, Distribution, and Sale of Drugs," Department of Justice, Ottawa, app. Q, p. 145.

Dr. COMANOR. That is correct.

SenatorNELSON. And then the sales price is different in that you add your profits and your advertising and your promotion, is that what you are saying?

Dr. COMANOR. That is what I am saying.

These statistics imply, then, that production costs are approximately one-third of the total selling price.

Senator NELSON. Where do you get that figure?

Dr. COMANOR. We are dealing with the same expenditures on quality control, and these outlays are taken as a fraction of two different quantities. One is total production costs and the other is total revenues. One percentage is a third that of the other.

Senator NELSON. You have several factors in here that I do not understand. You are talking in one case about quality controls. Quality control is not the same as production costs?

Dr. COMANOR. Oh, no; in fact, quality control is only 3.6 percent of total production costs.

Senator NELSON. And 1.2 percent of sales cost?

Dr. COMANOR. 1.2 percent of total sales revenues.

Senator NELSON. OK.

Dr. COMANOR. These statistics suggest that relatively small firms should be able to invest in quality control facilities and as can be seen, these expenditures are considerably less costly to the smaller firm or new entrant than are heavy advertising and promotional efforts.

Senator NELSON. As to these statistics that you got from the U.S. subsidiaries of manufacturers in Canada, does the Canadian Government require a breakdown of these various costs?

Dr. COMANOR. I do not think this breakdown is reported every year. A report published by the Restrictive Trade Practices Commission in Canada included these statistics.

Senator NELSON. When was that report published?

Dr. COMANOR. I am afraid I do not have the date here. Within the last few years. I have a copy of it.

Senator NELSON. Is it a large report?

Dr. COMANOR. It is quite a large report. It must be 500 or 600 pages.

Senator NELSON. In this report, do they have the statistics on the costs of production of a drug and the cost of quality control and the sales and profits—all of these statistics?

Dr. COMANOR. Not per drug, but for the entire company.

Senator NELSON. All together?

Dr. COMANOR. Or it may not even have been exported for individual companies but rather for the total of these 21 firms, only some of which were U.S. subsidiaries.

Senator NELSON. Did they not have any breakdown company by company? They must have gotten the figures from some place.

Dr. COMANOR. They may not have published them.

Senator NELSON. In order to get the composite, you have to have the individual figures.

Dr. COMANOR. They may not have published them. I am not certain about this point. I could find the answer for you.

Senator NELSON. If you would give use the name of the report, I will have the staff look into it.

Dr. COMANOR. Footnote No. 11 provides the reference to the report.

Senator NELSON. All right.

Dr. COMANOR. While more work needs to be done in gathering and analyzing statistical evidence on this problem, there is no evidence which indicates that only branded drugs produced by large firms are of a high quality.

THE CRITICAL ROLE OF INFORMATION

The analysis above suggests that advertising and other forms of promotion have an important effect on competition in the pharmaceutical industry. These outlays permit the leading firms in the industry to enjoy a considerable freedom from competitive contraints, and they have been translated into very high rates of return.

Furthermore, the analysis suggests that advertising and other forms of promotion have been effective to a large extent, because of the high degree of uncertainty about the relative merits of competing products which is present in the minds of the physician who serves as the purchasing agent for the final consumer.

On these grounds, there is much to be said for providing the physician with an impartial source of information which will relieve him of his current reliance on information provided by the drug companies.

Senator NELSON. I think this is a very good point. The fact is that a good part of the thrust of the industry's position is that you can't rely upon a drug unless it is produced by an established, well-known brand name company and that you are taking a risk with drugs from other companies. They end up convincing the physician of that. He has become convinced that there is not any such thing as generic equivalency or that it is a very serious problem. Further, there is not any available compendium of drugs that the doctor is prepared to trust, and that can assure him that he can prescribe a generic drug and trust its reliability. You are suggesting, I take it, that there ought to be a reliable, independent source so that the doctor could select drugs that are asserted by an independent source to be equivalent so that he can prescribe a drug that gives his patient a better price break.

Is that what you are saying?

Dr. COMANOR. That is correct. It seems to me that advertising is effective primarily because of the uncertainty which exists in the mind of the purchasing agent, which in this case is the physician. And if these expenditures were not so effective, we might expect a lower amount to be spent. What is important is to reduce the effectiveness of heavy advertising outlays by providing the doctor with an independent and reliable source of information.

This is not to say that the drug companies do not provide accurate information, although it may sometimes be less than complete, but rather that information from an independent source would have an element of impartiality which salesmen or printed advertisements rarely achieve.

Senator NELSON. If it is less than complete, it is not accurate, is it?

Dr. COMANOR. I agree that the drug companies do not always provide full information about the relative merits of competing drugs.

Senator NELSON. If it omits certain information to the doctor, then it is not accurate, is it?

Dr. COMANOR. That is correct.

Senator NELSON. But what you are saying is that they do not necessarily say anything incorrect, but if they omit something, the information is not accurate?

Dr. COMANOR. That is right. There are certain types of information which it is very important for the doctor to receive and which he rarely receives under the current system. He rarely receives a statement that two drugs are the same, because the incentive for the drug company is the opposite—to differentiate their product from those of their rivals. One piece of information which it is important for him to receive is that two drugs have the same or similar therapeutic effects, if that in fact happens to be the case.

Senator NELSON. In the majority of cases, no company is going to advertise that several other companies have a drug that is equivalent to its own.

Dr. COMANOR. That is right.

Senator NELSON. So it has to come from somebody else.

Dr. COMANOR. There are certain types of information which the doctor simply does not receive from advertisements. Therefore, he is left in the dark.

And also, and this seems to me to be equally important, it would be open to the smaller firm and new entrant as well as the large company.

Where various firms produce the same chemical compound, the physician should be informed either that all products are equally good, in terms of purity, strength, or some other characteristic, or that one brand is to be preferred to another.

And where different compounds compete, impartial information would lead to improved medical practice as well as reduce the significance of advertising and promotion.

If a small firm or new entrant introduced a beneficial new drug, it would compete on the basis of therapeutic properties rather than advertising claims. There is no evidence, moreover, that smaller firms are at a disadvantage in research as compared with advertising and promotion.[12] On many grounds, then, there seems to be a considerable need to provide the Nation's doctors with greater information about new drugs which is independent of the messages they receive from the drug companies.

Mr. GORDON. Dr. Comanor, in the footnote on page 16, you state "The available evidence suggests, moreover, that small firms have relatively more efficient research facilities than their larger rivals."

What evidence do you have to show this?

Dr. COMANOR. In one of the papers which I have published, I report on a statistical analysis of the relationship between research input and research output.

Senator NELSON. Your analysis was comparing the products resulting from the research?

---

[12] The available evidence suggests, moreover, that small firms have relatively more efficient research facilities than their larger rivals. William S. Comanor, "Research and Technical Change in the Pharmaceutical Industry," *Review of Economics and Statistics*, May 1965.

Dr. COMANOR. That is right. I carried out a study of a large number of pharmaceutical firms and I measured the relationship between research input and output. I measured the output of a research establishment by the number of new products introduced by the firm since research in the pharmaceutical industry is directed largely toward new products rather than new processes. The number of new products was multiplied by its sales during the first 2 years following introduction. To make these computations, I obtained access to the Gosselin data which provides sales by product for all products sold in the country. Research input was measured by the number of research personnel employed by the individual pharmaceutical firm, and the relationship between input and output was then estimated.

Senator NELSON. You did not make, I take it, any distinction between the efficiency of the research of one company in developing a new drug versus another company in producing a new drug? You did not make a distinction between that and the research which produced a product differentiation, did you?

Dr. COMANOR. No, I did not.

Senator NELSON. So whatever they produced, either as a product differentiation or a new drug, is included in the statistics?

Dr. COMANOR. I used two measures of new products, but the most important included only new single chemical entities. I was, therefore, dealing with products which were new in the sense that they were different chemical compounds which had previously been introduced in the U.S. market.

While it is true, of course, that some of these embodied a greater therapeutic advance than others, and it may be the case that some embodied no or very little therapeutic advance, I did not feel capable of deciding this question. I simply took all new single chemical entities, multiplied by their sales in the first 2 years following introduction, to measure the output of research.

When I carried out the statistical analysis between research input and research output, which is reported in one of the papers that I have published on the industry, there were a number of findings:

First, there was a fairly sustained and direct association between input and output. In other words, research is not undertaken in this industry simply in the hope of some payoff in the future, but rather there is this fairly sustained and systematic relationship between the amount of research undertaken and the number of new products introduced by the individual firm.

Senator NELSON. When you say you are comparing the relatively small firms versus their larger rivals, what standard did you use to distinguish a relatively small firm and a large firm?

Dr. COMANOR. I measured the size of firm by the level of pharmaceutical sales.

Senator NELSON. Yes; what was the cutoff point?

Dr. COMANOR. There was no cutoff. It was undertaken not on a discrete but on a continuous basis. I examined the entire relationship between small firms and large firms without any arbitrary cutoff point to distinguish large from small.

Senator NELSON. So some of of them were relatively small?

Dr. COMANOR. Yes, and I had firms of all sizes in my sample.

I found that economies of scale exist in very small firms, but at the same time, diseconomies of scale exist in large firms and even in research carried out in moderate-sized firms. From these statistical results there appear to be no advantages from increasing the size of the firm's research effort beyond 50 to 75 professionals. At the same time, we can note that most of the leading firms in this industry have research labs which are far larger than this size. The cutoff point, which was not determined arbitrarily but resulted from the statistical analysis, was that economies of scale appear to exist up to the level of 50 to 75 professional personnel, but become diseconomies after that level.

The position stated above is identical to that taken earlier by the Assistant Attorney General in charge of the Antitrust Division, Mr. Donald F. Turner, who in a recent speech discussed the competitive effects of advertising expenditures generally throughout the economy.

Mr. Turner emphasized:

> It is the extent of uncertainty about the relative merits of competing products which contributes to the large effect of advertising, and this suggests that Government policies be directed toward neutral vehicles of information which tend to deal directly with the uncertainty.
>
> He goes on to mention the Medical Letter, which is an important source of "technical information about the therapeutic value of new drugs", and concludes that "there is much to be said for providing Government funds to the organization which publishes the Medical Letter so that its publication may be supplied free to all doctors. In addition, the latter could be expanded to ensure that doctors receive their first information about a new drug from this source rather than from the lips of a detail man." [13]

Mr. GORDON. Is it your opinion that this measure could offset the huge expenditures by the drug industry on advertising and promotion?

Dr. COMANOR. I think that this type of policy, while it may not offset fully the impact of advertising and promotion on product differentiation and thereby on price competition, would have an important effect. I would expect that it would reduce significantly the effectiveness of these outlays and think we could expect a higher degree of price competition with this information provided to all doctors than we find currently.

I agree fully with Mr. Turner's position that an impartial and independent source of information is needed—that it would be an important step forward, both in improving the quality and standard of medical practice throughout the country, and in reducing the anticompetitive consequences of heavy advertising.

There are a number of alternative routes to achieving this objective. One might be to provide Federal support to the nonprofit organization which publishes the medical letter. Not only would this permit a more comprehensive study of drugs and drug claims than is possible on the basis of available facilities, but also it would allow an expanded Medical Letter to be sent free to all phyisicians throughout the country.

Mr. GROSSMAN. Dr. Comanor, may I ask a question?

Several of the economists who have appeared here have talked about the fact that we should get rid of the brand name, trade name, and just use the generic name.

Do you feel that way as well?

Dr. COMANOR. There are a number of routes to achieving the same objective. One route, in some sense an easier route, would be to provide

[13] Turner, p. 2054, supra.

information to the doctor which would reduce the impact of trade names and heavy advertising outlays.

Another route, which is perhaps a stronger step, would be to eliminate all trade names altogether. I certainly feel that further information is needed now regardless of whether or not trade names are eliminated. At the same time, it seems to me that trade names in this industry have become a very important source of limiting the entry of new firms into the relevant therapeutic markets and restricting the development of price competition. Therefore, although I have not emphasized this problem in my statement, I would certainly subscribe to this position.

Mr. GROSSMAN. If we did away with trade names, would it not be likely that we would have a generic name (Schering), generic name (Upjohn), and the company would promote their name? They would find a way to get around it?

Dr. COMANOR. There is no question that the drug companies would aim to achieve just the end which you suggest. One might argue that their ability to differentiate their products on this basis would be less than it is currently, and as a result, we might find a higher degree of price competition. This is not an all-or-nothing situation; it is clearly a matter of degree. While some measure of product differentiation would remain, and might even be desirable at a small level, it would not lead to the high prices which currently exist and the very high rates of return.

Mr. GROSSMAN. Do you think doctors should be given some information as to variable pricing such as the formulary?

Dr. COMANOR. I would like to see the price of the product included in the Medical Letter or any sources of information provided to doctors.

Mr. GROSSMAN. Once that is done, the doctor should be able to make the choice? So long as he is given this information, he should be able to make a free choice?

Dr. COMANOR. That is a very difficult question. Although price information should be provided to physicians, I do not think I would go so far as to say that in every case, only the doctors should decide among generically equivalent drugs—that the patient should not be permitted to achieve generic equivalency if this were desired. I am not prepared to say that there should be no option to the patient on generic equivalency.

Senator NELSON. The patient does not have the option on generic equivalency.

Dr. COMANOR. No.

Senator NELSON. The doctor does.

Dr. COMANOR. The question I think we are really getting at is whether or not the pharmacist should be permitted to substitute among generically equivalent drugs. I find that it is a very difficult problem. I am not prepared to say definitely that providing since information to the physician is sufficient.

Mr. GROSSMAN. If we list various drugs under a generic heading, be they trade names or generic names, and we also list the prices next to them, will we have gone far enough?

In other words, we will be giving the doctor quite a bit of information, more than he has now.

Is this going far enough?

Dr. COMANOR. I think that is an important step forward. Whether or not that would be sufficient is a question which is very difficult and I am not prepared to say definitely that it is.

Mr. GROSSMAN. But it would be a step forward?

Dr. COMANOR. I think so.

An alternative proposal would be to create a new institute in the National Institutes of Health which would be responsible for evaluating existing information on the therapeutic properties for new drugs, carrying out their own studies where necessary, and also for providing the accumulated information freely to all physicians.

In this manner also, an independent source of information about both current and new drugs would become available. By whatever route this is achieved, however, it seems clear that the relatively small level of expenditures required for these purposes would pay for itself many times in terms of limiting the effects of heavy advertising outlays and in terms of lower prices for prescription drugs.

While a new source of information will not solve all of the competitive problems which exist in the pharmaceutical industry, and restrictions on competition will remain because of the current operation of the patent system; because of the practice of trade name prescribing; and because of the problem of drug substitution or the lack thereof at the level of the pharmacist, still Federal action in this direction should make an important contribution toward improved medical practice as well as toward lower prices for prescription drugs.

Mr. GROSSMAN. Independent economists have previously recommended compulsory licensing as one means of making the industry more competitive.

Would compulsory licensing lead to lower prices?

Dr. COMANOR. The patent system currently operates to provide a form of chemical product differentiation in the sense that competition exists on the basis of different chemical entities. While it is true that different products compete with one another, the patent system increases the extent to which competition is based on different chemical entities. And this always provides an important element of uncertainty to the physician. He is never quite sure whether one product is better than another.

For this reason, it increases the effectiveness of advertising. It permits advertising to have a major impact on prescribing habits, on the entry of new firms, and on price competition.

With compulsory licensing, we might expect to find new firms enter the industry, competition take place on the basis of the same chemical compound, and here, advertising would probably be less effective than it is currently, especially if further information were provided to physicians. And we might expect to find a higher degree of price competition, and as a result, lower prices.

Mr. GROSSMAN. Now, would that lead to less research by the drug companies?

Dr. COMANOR. With compulsory licensing, we would expect to find different firms competing on the basis of the same chemical com-

pound. This would tend to reduce, as I have noted, the effectiveness of heavy advertising. And we might expect to find more vigorous competition and lower prices.

An example of the effect of patent protection is the major difference in price behavior between penicillin, where price competition has been vigorous, and the broad spectrum antibiotics, which for a long time were patent protected, when there was little price competition.

At the same time, we have to expect that with lower rewards from research, it is likely that less research will be undertaken, and it is important to recognize that there is in fact a tradeoff between industrial research and lower prices.

However, it seems to me that we pay a high price for industry research in terms of the high prices and profits which exist in this industry. A major question for public policy concerns what we get for this price. To answer this question, we have to think about the character or payoff which society realizes from industrial research. We have to consider the value of industrial research.

Industrial research is carried on within the context of strong product rivalry. This rivalry influences significantly the type of work which is in fact carried out. We can note that there is a large effort to invent around existing patents, to find a new product which has similar therapeutic properties to those already on the market. And this, of course, is what we mean by molecule manipulations, which are purported to have the same therapeutic effect as existing products.

Also, if we examine the research efforts and output of the industry, there appears to be a considerable effort to create new drug combinations, combining different active ingredients with one another, and to introduce new dosage forms.

As a result, we do not expect the major discoveries to be made within the industry precisely because of the character of the research effort which in fact is undertaken. Still, it seems to me that the minor advances which result from the research laboratories within the pharmaceutical industry, the modifications and variations which result, are important elements in the process of developing a new technology.

The second major emphasis of industry research is the process by which new knowledge is translated into marketable products. We know that many problems remain even after a basic discovery has been made. It has been estimated, in fact, that 50 percent of research and development expenditures within the pharmaceutical industry goes to support the work of synthesizing, purifying, modifying and preparing substances for testing. And we know that considerable testing is required before a drug may be released to the public.

In addition, there is also frequently a considerable gap between laboratory synthesis and large-scale manufacturing which is also carried out within the research laboratories of the pharmaceutical industry.

Although this might be considered development work, it is still quite necessary. And we have only to note that although penicillin was discovered originally in 1928, it was not used effectively until some years later, until the Second World War, 1943.

This time lag is likely to be much shorter because of the current facilities of the drug industry.

To evaluate industry research and development, however, we have to place it within the context of the total medical research and development effort in the United States. In 1960, the industry provided about 30 percent of total research and development outlays. At the same time, it provided only 12 percent of the Ph. D.'s involved in medical research and employed only about 5 percent of the M.D.'s in medical research. This indicates, thus, that the character of research in the industry is likely to be far different, at a lower level of scientific caliber, than the medical research which is carried on outside of the pharmaceutical industry.

It seems to me, therefore, that the industry research effort cannot be evaluated or judged by the same standards as nonindustry research. It seems to me that the industry research effort should be considered as a complimentary rather than as a competitive or alternative type of activity. There appear to be extensive resources which are devoted to fairly routine sets of activities.

When evaluated on its own terms, on a different basis than might be appropriate for evaluating academic or Government medical research, it seems to me that the industry effort, by and large, does a good job. It has, in fact, been responsible for the large number of modifications and improvements of existing drugs. It has also been responsible for accelerating the process of testing and developing new drugs. And I would argue that both of these provide important benefits to the development of medical technology and medical practice in the country.

At the same time, the question remains as to whether the "price" that society pays in terms of the opportunity costs for these results is too high. If we examine the proposal of compulsory licensing after a short period of time, such as the Kefauver suggestion of 3 years, it is clear that this proposal would reduce the incentive to undertake research.

The industry certainly would have less incentive to undertake purely duplicative research in order to enter a profitable therapeutic market.

Senator NELSON. If they had less incentive to do purely duplicative research, that would be a net benefit to society; would it not?

Dr. COMANOR. It certainly would not be any loss to society.

Senator NELSON. And they would tend to do more pure research to develop new products and that would be beneficial; would it not?

Dr. COMANOR. I think that is right. At the same time——

Senator NELSON. What I am really getting at is that we have had testimony that much of the research is directed solely at product differentiation, with the production of a product that does the same thing another drug does, and in some instances not even as good a job as another drug. I understand that the argument the industry will make is that as a result of this kind of research and molecular manipulation, we sometimes get an improved product. You recognize that, too.

But if it did eliminate just purely duplicative research, the result of which is to produce just another product doing the same thing as a drug already on the market, then it would seem to me that it would be beneficial; would it not?

Dr. COMANOR. Yes; that is correct.

With compulsory licensing, it seems that the incentive to undertake purely duplicative research would be clearly much less. At the same time, the incentive to undertake expensive basic research projects would also be much less. Basic research work is likely to be much more costly than applied work, as well as more risky, and if the returns from this work would be lessened through compulsory licensing, the industry would probably carry out much less pure fundamental research than it does now.

Senator NELSON. I do not follow the argument, because if they produced a new product that is useful, then every doctor who has a use for that type of drug is going to prescribe it. They then could make their profit over some period of time and when they are forced to license other firms they would receive a royalty from what is sold all over the world.

Why is that not an incentive?

Dr. COMANOR. That is an incentive, but the important question is not simply whether an incentive exists but how much. And clearly with compulsory licensing, the prospective gains would be less than they are currently, because currently, they can charge high prices throughout. With lower prospective gains, we might expect to find less research undertaken of a very costly nature, and basic research is probably the most costly type of work which is undertaken.

At the same time, it seems to me that with regard to basic research, it is difficult to argue that private firms have a comparative advantage relative to Government, foundations, and universities. Most basic research in the medical area is performed outside of the industry now, and even if the industry carried on no basic research, it is not clear that this charge would involve heavy social costs. Basic research, as many commentators have noted, depends on an atmosphere of freedom of inquiry, which may not be best achieved within the confines of a profitmaking organization.

It seems to me, then, that the comparative advantages of basic research probably lies in universities and Government laboratories. While the magnitude of industry basic research would decline, this is precisely an area where the industry does not seem to have a comparative advantage. At the same time, there would still be an incentive for the firm to differentiate its products from those of its rivals and to create entry barriers. Furthermore, it would continue to earn high returns during the first 3 years, and receive patent royalties thereafter.

The industry would be likely to emphasize the area where immediate payoff is greatest. They would emphasize modifications which represent relatively small advancements in medical practice.

They would also, it seems to me, be expected to emphasize the development and testing of new drugs which were originally discovered outside of industry laboratories. Before a new drug is introduced, not only must the basic discovery be made but also considerable scientific and technical resources are required to test it under many different conditions. This is precisely the work in which the industry appears to have a comparative advantage. This is the sort of work in which the payoff specifically from industry research is likely to be greatest, and we might expect to see it continued even with compulsory licensing.

As you note, high gains would be obtained during the original 3-year period, and in addition, of course, gains from patent royalties would remain.

There are important areas where drug companies probably do have a comparative advantage in terms of the type or character of research, but also it seems likely that this type of research effort would continue even with compulsory licensing. Therefore, I would certainly subscribe to the view that compulsory licensing, after a period of time at reasonable royalty rates, is a desirable policy change and should be enacted.

While it is true that the magnitude of research and development in the pharmaceutical industry would probably decline, it seems likely that projects of limited medical value and lowest industry comparative advantage would be eliminated first.

Senator NELSON. You are talking about pure research as contrasted, say, with applied research.

The industry does not do very much pure research now, does it?

Dr. COMANOR. There is, of course, considerable controversy over precisely how much pure research drug firms currently undertake.

Senator NELSON. It is product-directed; is it not?

Dr. COMANOR. It certainly is.

Senator NELSON. Is it not true, by most definitions, that there is not very much pure research done by industry as a whole in this country as compared to applied research?

Dr. COMANOR. I think the correct statistic, throughout the entire economy is something on the order of 10 percent, but I would have to check to be completely accurate. But however much basic research is done now, I would expect that less would be done with compulsory licensing.

At the same time, I am not unhappy with that result, because as I have noted, this is not an area where the industry appears to enjoy a comparative advantage relative to nonindustry research.

Therefore, I would suspect that with compulsory licensing, the magnitude of pure research would decline, but research with the highest industry comparative advantage would continue. I think it inappropriate to argue that the more research the better in this type of profit-oriented context, and it may well be that the present research effort carried out in the industry is sufficient or even exceeds that which is needed for its major research functions and responsibilities.

Senator NELSON. What percentage of sales does the research effort represent now in the industry?

Dr. COMANOR. Different sources provide different estimates. From the statistics I have seen, it is on the order of 6 to 8 percent.

Senator NELSON. Of sales?

Dr. COMANOR. Of sales.

Therefore, the lower research which would result from compulsory licensing seems to me to be a small price to pay for a more competitive determination of prices and profits in this industry.

Mr. GORDON. Dr. Comanor, the relevant market of the drug industry, as I understand it, is not the industry as a whole but each individual, therapeutic category. Would you say that is correct?

Dr. COMANOR. When economists define the limits of a market, their aim is to determine the area, whether in terms of geographic limits or product class limits, in which prices are set, and this generally depends on the area in which different products may be considered as highly substitutable. And if we examine the question of substitutability in the pharmaceutical industry, it is clear that a diuretic, which is used for one type of ailment, cannot be substituted for an antibiotic, which has a very different use. Given the very low degree of substitutability among therapeutic classes, I think most economists would agree that the relevant markets here, which define the limits within which prices are determined, is the therapeutic class rather than the industry as a whole.

Mr. GORDON. What do you have to say about the concentration of output in each of these therapeutic categories?

Dr. COMANOR. If one examines therapeutic markets, the concentration ratios are very high. This is precisely because effective entry barriers have been created. If advertising and research had not been able to create effective product differentiation, we would expect high prices to attract new firms into the market. That this has not occurred is due to the factors which I have discussed.

Mr. GORDON. You have stated that price competition is virtually absent from this industry.

What do you have to say about the drive to reduce costs and produce drugs more efficiently?

Dr. COMANOR. Since production costs account for such a small proportion of total costs in this industry, on the order of a quarter to a third, there is clearly relatively little to be gained by the firm from putting a considerable effort in producing drugs more efficiently. Moreover, lower production costs would have little competitive effect. With little to be gained, it is not surprising to find a relatively small effort devoted to reduce production costs. While I am afraid that no hard statistical evidence is available on this question, it is likely that the drive to reduce costs is much less than would be the case if price competition were more effective and more vigorous.

Mr. GORDON. So actually, we have a situation, do we not, where you have not only product differentiation but also a highly concentrated market? What would be the effect of both of these on the possibility of price competition?

Dr. COMANOR. We observe that there is relatively little price competition among highly differentiated products, although there does seem to be some degree of price competition among those products which are not effectively differentiated. For example, we still see effective price competition for penicillin and streptomycin.

But, as we can tell from the high overall profit rates, these must account for a relatively small share of the total number of products produced. During the period of the Kefauver committee investigations, it was reported that—I think this was for 1957—the profit rate in the pharmaceutical industry was double or more than double the average of all manufacturing.

Mr. GORDON. Even if you did not have product differentiation, the high concentration itself would tend to bar the possibility of competition in this industry?

Dr. COMANOR. That is true, but high concentration and entry barriers are related to one another. The primary reason for high concentration in relevant therapeutic markets is precisely because of the entry barriers created by product differentiation. So it is difficult to distinguish the competitive effects of one from the other.

Mr. GORDON. Now, given the type of research conducted by the industry, and I want to highlight this point which you have already brought up, would it be fair to say that the industry is heavily dependent on new scientific knowledge developed outside the industry?

Dr. COMANOR. I think it is clear that much of a research effort in the industry makes use of new scientific advancements which come about in nonindustry laboratories. At the same time creating new drugs out of new scientific knowledge is precisely what might be considered the major research responsibility for the industry.

So I certainly would not be critical of this fact.

It seems to me this is precisely what society should expect and demand from the industry, that it will utilize the research efforts from outside, where the majority of medical research is in fact undertaken, and translate important discoveries into marketable products.

So while much of the industry research effort is devoted to this type of work, I feel that this is precisely what is desirable or useful from the point of view of society.

At the same time, this type of work would probably be carried on even under a regime of compulsory licensing after 3 years, and therefore, I think we have much to be gained from adopting this policy.

Senator NELSON. Thank you very much, Dr. Comanor. Your testimony was a very useful and valuable contribution to the record. We appreciate your taking the time to come here and testify today.

Dr. COMANOR. Thank you very much.

(The supplemental information submitted by Dr. Comanor follows:)

[From "Economica," November 1964]

RESEARCH AND COMPETITIVE PRODUCT DIFFERENTIATION IN THE PHARMACEUTICAL INDUSTRY IN THE UNITED STATES [1]

(By William S. Comanor)

During the past two decades expenditures in the United States on research and development have expanded greatly. From outlays of the order of a half-billion dollars at the start of the Second World War, they have grown to well over $10 billion. The rapid growth of an "industry of discovery" [2] has given rise to a large number of questions concerned with the rôle and function of industrial research. Why do firms spend large amounts on research and development? What is the relationship between research and market structure? What has been the impact of public policy, in the form of the patent system, on these expenditures?

In dealing with questions of this sort it is necessary to stress the pattern of relationships among firms within which research and development is undertaken. In an activity characterized by high degrees of uncertainty, we should not expect

[1] This article is based on the author's unpublished Ph.D. dissertation, *The Economics of Research and Development in the Pharmaceutical Industry*, Department of Economics, Harvard University, June 1963. I would like to acknowledge the assistance of J. W. Markham and R. B. Heflebower, who acted as supervisors of the original study, and also of E. T. Penrose for a number of helpful suggestions in the writing of this article. The study was undertaken with the aid of a research grant from the Science and Public Policy Programme at Harvard, and I am indebted to the Programme for this support.

[2] John T. Dunlop, "Introduction: Problems and Potentials", in John T. Dunlop, ed., *Automation and Technological Change*, p. 2.

firms to invest so heavily in the absence of specific pressures. These pressures arise most frequently from oligopolistic market structures in which the need for product differentiation has been recognized and where rivalry has been limited to variables other than price. This article puts forward an analytical framework within which these questions may be considered, and also attempts to provide some answers dealing with the experience of the pharmaceutical industry in the United States.

THE GROWTH OF RESEARCH AND DEVELOPMENT

The spectacular growth of the pharmaceutical industry in the postwar period has been founded upon a number of important medical advances and has been associated with some radical changes in the structural and behavioral characteristics of the industry.[3] Prior to 1937 and the introduction of the early sulpha drugs the industry was composed largely of long established firms producing relatively standardized commodities. Unlike as at present, a large proportion of medical prescriptions were compounded by the pharmacist.[4] Barriers to entry were low and a high degree of competition prevailed.

With the introduction of penicillin and streptomycin during and immediately after the Second World War, the nature of the industry changed. As neither of the products was protected by patents,[5] the rapid growth of demand resulting from their introduction was accompanied by the entry of many new suppliers and by the development of active price competition. The price of a standard form of penicillin dropped from $20 for 100,000 units in 1943 to 4½ cents in 1950[6]; the price of streptomycin behaved similarly.[7] Excess capacity was created, and it may well be that prices fell to the vicinity of short-run marginal costs.

Following this experience, it was clear to the leading firms that their profits in the future would depend on the development of more protected market positions to be achieved through some form of product differentiation. This would provide the producer with substantial control over the prices of his products as well as act as a significant barrier to entry into the relevant therapeutic market. With the trend in the new technology towards the compounding of prescriptions by the manufacturer rather than the pharmacist, this development seemed all the more promising. Although increased selling expenditures would be necessary, these were not likely to be sufficient to establish effective differentiation, for the industry faced a relatively informed consuming public in the medical profession, and it would be more difficult to differentiate between products consisting of identical chemical substances. Only through the extensive introduction of new products could significant differentiation be achieved. This would enable individual firms to emphasize in their selling activities the improved quality of new drugs. Moreover, if each of the leading firms followed this course of action, competition in the industry would largely cease to be founded on the sale of standardized commodities.[8]

The accelerated growth of pharmaceutical research and development expenditures during the post-war years[9] was due largely to the desire to promote the

---

[3] In this study the pharmaceutical industry is defined to include those firms which produce ethical drugs, as opposed to proprietaries, and which distribute these products in dosage-form. Pharmaceuticals are, thereby, marketed and sold only through the medical profession and require, for the most part, a written medical prescription.

[4] In 1961 prescriptions compounded by pharmacists accounted for only 3.6 per cent of all prescriptions in the United States.

[5] While a patent was issued to the inventor of streptomycin, Dr. Selman A. Waksman, it was assigned to the Rutgers Research and Endowment Foundation, and was licensed on a relatively liberal basis.

[6] Federal Trade Commission, *Economic Report on Antibiotics Manufacture*, 1958, p. 230.

[7] See United States Senate, Subcommittee on Antitrust and Monopoly, *Report of the Study on Administered Prices in the Drug Industry*, 87th Congress, 1st Session, 1961, pp. 81–8.

[8] In this context it is instructive to note the statement made by an industry spokesman in 1950. When considering the industry's future, he declared: "From a profit point of view . . . the only realistic solution of this problem lies in the development of *new and exclusive antibiotic specialties*. This . . . is an exceedingly costly and vigorous alternative; nevertheless, it is the avenue of approach being most extensively explored by certain antibiotic houses today. This is the approach being followed by Pfizer". Statement by John McKean, President of Chas. Pfizer & Co., quoted in Senate Report, *op. cit.*, pp. 130–31 (italics added).

[9] Between 1951 and 1960, research and development expenditures in the pharmaceutical industry grew from $60.4 million to $206.5 million, an increase of over 200 per cent. (Pharmaceutical Manufacturers' Association, *Yearbook, 1961–62*, p. 168.) These figures are measured in current dollars and are derived from surveys of P.M.A. member firms.

introduction of new products and the accompanying extension of product differentiation. Research laboratories may, in fact, be regarded as a cost of differentiation. Facilities were designed for the discovery, development, and testing of new drugs, and as a result, new products were introduced at a rapid pace. While completely new chemical entities were the most desirable type of new product, since these would provide the highest degree of differentiation, other forms were also useful for this purpose, and in addition frequently had the advantage of requiring less research input per product. Among the latter group were new combinations of drugs already on the market, and new dosage-forms of previously introduced products. During the decade ending in 1960, over 3,800 new products and dosage-forms were introduced into U.S. pharmaceutical markets, or an average of more than 380 per year. Of these slightly more than 11 per cent. of the total were new single chemical entities.[10]

THE IMPACT OF COMPETITIVE PRESSURES

In examining the impact of competition on the growth of research, it is useful to consider the rôles played by certain structural characteristics of the industry. The most significant of these is the co-existence of a number of fairly large firms with a swarm of smaller ones, which, however, account for a substantial proportion of the total output of the industry. Relevant data are presented in Table 1. As may be observed, the industry overall is moderately concentrated, with a significant fringe of smaller firms.[11] Some further important characteristics concern the relatively low proportion of variable costs in total costs—in the larger firms the proportion tends to be less than half[12]—and a generally low price elasticity of demand for pharmaceuticals.[13] In a siaution of this sort, not only are competitive price declines likely to be large but also firms could easily be compelled to price below average total costs. As a result, there would be heavy pressures on profits if rivalry were allowed to take the form of price competition.

TABLE 1.—PHARMACEUTICAL INDUSTRY: CONCENTRATION RATIOS, 1958

	Sales	
	Amount (millions)	Percent
Largest 3 companies	$384	20
Largest 7 companies	800	41
Largest 15 companies	1,190	61
All companies	1,945	100

Sources: Pharmaceutical Manufacturers' Association, "Domestic Human Ethical Drug Business Survey," conducted by Arthur Andersen & Co., 1959, mimeographed. Pharmaceutical Manufacturers' Association, "Yearbook, 1961-62," p. 170.

The data for the largest 15 firms are derived from a survey of sales in 1958 of 50 pharmaceutical firms, all members of the Pharmaceutical Manufacturers' Association. Given the nature of the survey, it can be assumed that the 50 companies included the largest 15 in the industry. The figure for industry sales refers only to P.M.A. membership; it has been maintained by the association that member firms account for over 95 percent of total industry output.

The interaction of these two sets of factors has been significant in determining industry behaviour and performance. Where smaller firms are relatively important, one might expect to find a high degree of price competition. That this result has not occurred can be attributed to the post-war strategy of research and product differentiation pursued competitively by the leading firms in the industry. Rivalry has been restricted largely to areas other than price. A high degree of price stability on existing products has been maintained, and new products have been priced, for the most part, to compete with older ones in the same

[10] Paul de Haen, *New Product Survey*, 1960.

[11] The pharmaceutical industry falls easily into Kaysen and Turner's Type II oligopoly. According to their definition, a Type I oligopoly is one in which the largest eight firms account for at least 50 per cent of industry output while the largest twenty comprise at least 75 per cent. A Type II oligopoly, on the other hand, is one in which the largest eight firms account for at least 33 per cent. of industry output. See Carl Kaysen and Donald F. Turner, *Antitrust Policy*, p. 27.

[12] In a sample of twenty-two largest pharmaceutical firms, and with regard to 1958 drug operations only, costs of goods were only 43 per cent. of total costs. See Senate Report, *op. cit.*, p. 31.

[13] See *The Economics of Research and Development in the Pharmaceutical Industry*, pp. 58–60. A similar conclusion was reached in the Senate Report, *op. cit.*, p. 3.

therapeutic class,[14] At the same time, however, a vigorous form of rivalry has developed, which has taken the form of what Dr. Penrose has termed "competition in creativity".[15] A policy of competitive innovation seems to have been formulated by the individual firms, and research laboratories have been established as a vehicle of this policy. Profits have come to depend primarily on the firm's position in the innovative race, and the cost of an organized and efficient research effort may be regarded as a necessary expense of entering the race. Once some firms had taken this position, moreover, others were forced to follow in order to maintain their market positions relative to their rivals.

In this manner the growth of pharmaceutical research establishments has been stimulated to a considerable extent by competitive pressures. Not only has the latent price competition, arising from these pressures, impressed upon the larger firms the need to achieve effective product differentiation, but also a highly competitive form of research and development has been promoted. Competitive factors have influenced the direction and scope as well as the size of research establishments.

PRODUCT COMPETITION WITHIN THERAPEUTIC MARKETS

The expansion of pharmaceutical research facilities and the concomitant increase in the introduction of new products have given rise to a rapid rate of product obsolescence. New products have competed with older ones, and the latter have been rapidly replaced by the former. By 1960 only 31 per cent of sales accrued from products introduced before 1951.[16] As may be seen in Table 2, the sales of products introduced in a given year normally reached their peak (as a proportion of total sales) in the second year following their introduction; after that they declined rapidly in relative importance. New products were introduced, reached their peak, and declined; and all within a relatively short period of time.[17]

TABLE 2.—SALES OF PHARMACEUTICAL PRODUCTS BY AGE OF PRODUCTS

Year of introduction of products	Percent of sales									
	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960
1960										4.6
1959									7.1	10.5
1958								4.7	7.2	7.3
1957							10.1	16.8	16.7	15.1
1956						5.8	9.0	8.4	7.7	6.6
1955					9.0	18.7	15.9	11.4	10.2	8.2
1954				7.4	13.7	12.8	10.8	10.2	8.8	8.1
1953			5.4	16.0	14.8	12.1	9.8	7.2	5.8	4.7
1952		4.6	9.5	7.7	6.1	4.8	4.4	4.0	3.1	2.9
1951	1.1	1.7	1.6	1.1	.8	.7	.6	.6	.6	.6
Before 1951	98.9	93.7	83.5	67.9	55.6	45.1	39.5	36.7	32.8	31.4
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Source: Arthur D. Little, Inc., "A Report on the Aspects of Concentration and Product Obsolescence in the Pharmaceutical Industry in the United States," p. 33.

The table is based on a sample of 33 firms, comprised, for the most part, of the larger firms in the industry. These firms account for approximately 80 percent of industry output.

[14] A similar view was expressed by the Senate Subcommittee, which reported: "When a new product is put on the market, the customary procedure is to introduce it at or very near the price charged for an existing drug used to treat the same general type of ailment". Senate Report, *op. cit.*, p. 98.

[15] E. T. Penrose, *The Theory of the Growth of the Firm*, p. 106. "In the United States . . . a kind of 'competition in creativity' has become a dominant motif in the pattern of competitive behavior in many industries, where consumers and producers alike are caught up in an almost compulsive obsession for that which is 'new'. In the extreme case the individual firm is forced constantly to remould its products—to create the 'new' and 'improved' either in performance or design. To a large extent the new products are superior in performance; to a considerable extent they are merely new and can be sold only if the consumer can be convinced that the 'newest' is the 'best'."

[16] Arthur D. Little, Inc., *A Report on the Aspects of Concentration and Product Obsolescence in the Pharmaceutical Industry in the United States*, p. 33.

[17] Somers and Somers report that the average life-span of a new drug is said to be between two and five years. Herman M. Somers and Anne R. Somers, *Doctors, Patients and Health Insurance*, p. 96.

Because of the segmented nature of demand, it is misleading to consider the market for pharmaceuticals as a single entity; the relevant market does not span the entire industry but, rather, is limited to individual therapeutic classifications.[18] And it is within these therapeutic classifications that the rapid rate of product introduction has developed into vigorous product competition. New products have replaced older ones, and in this process the ranking of the leading firms in particular therapeutic markets has changed frequently. In only nine out of twenty such markets did the same firm persist as the leading firm between 1951 and 1960; and in most cases the same five firms did not retain completely the five leading positions.[19] While past position tends to be an advantage in determining present market shares, it is not of crucial significance, and the measure of market control derived from a single product appears to be limited to a relatively short period of time.[20]

NEW PRODUCTS OR NEW PROCESSES

Research laboratories in the pharmaceutical industry have been concerned largely with the introduction of new products rather than with the development of new processes [21] for old products.[22] Given the nature of competition, this is what we should expect. Where product rivalry is high and the effective life of individual products is correspondingly short, firms are unwilling to invest large amounts towards reducing production costs. By the time new techniques have been developed, it is quite likely that demand for the product concerned will have dropped to a relatively low level. Even with regard to products which have fairly long lives, moreover, or to processes used to produce many products, it is unlikely that a large research effort is undertaken in the direction of reducing production costs. To the extent that rivalry is based on price, the introduction of new processes will place the firm at a competitive advantage by enabling it to price below its rivals without reducing its profit margins. When, however, price behaviour is largely non-competitive, a reduction of costs will increase profit margins without immediately affecting the firm's position relative to those of its rivals. While this may enable the firm to undertake larger selling or research expenditures, or to raise funds more advantageously in the capital market, the competitive effect will not be immediate or direct.

The introduction of new products affects the competitive position of the firm in a quite different manner. To the extent that rivalry takes the form of competition between products which are priced at the same or similar levels, the number and character of new products which are introduced directly affect the demand for the firm's output. If the firm introduces new products which do the job "better", then its output and total profits may be higher even if the costs of the new products are greater than those of their predecessors and profit margins correspondingly reduced. The firm that falls behind in the race to intro-

---

[18] Relatively high cross-elasticities of demand will be found when competing products treat the same or similar ailments or have similar therapeutic effects. Between therapeutic classifications, on the other hand, these elasticities will approach zero. While cases will remain in which a specific product is required to treat a certain condition, the gap in the chain of substitutes will appear at the boundaries of therapeutic classes. Although high cross-elasticities of supply will exist across these boundaries, we still find it appropriate for our purposes to define the relevant markets in terms of therapeutic classifications.

[19] Data describing the market shares of the leading five firms within twenty therapeutic markets for the years 1951 and 1960 are presented in Arthur D. Little, Inc., *op. cit.*, pp. 11–30.

[20] This conclusion is corroborated by McKie who reports the following: "In all cases it is clear that the period of 'dominance' of any one product is short—four or five years at most—and that a firm which fails to bring out improvements or new substitutes will find its share of the market rapidly passing to others". James W. McKie, "An Economic Analysis of the Position of American Home Products Corporation in the Ethical Drug Industry", in United States Senate, *Administered Prices in the Drug Industry*, Hearings before the Subcommittee on Antitrust and Monopoly, 86th Congress, 2nd Session, 1960, Part 17, p. 9957. The latter document will be cited as *Administered Price Hearings*.

[21] The distinction between products and processes cannot be based on characteristics of the good in question. Where intermediate products are concerned, a product for one firm may be a process for another. The distinction is actually relevant only within the firm. A product is a good which is offered for sale in the normal course of business. A process, on the other hand, is only infrequently offered for sale, and is used within the firm to facilitate production of other commodities.

[22] This emphasis was sufficiently great for a survey of research budgets of pharmaceutical firms to allocate total outlays among four categories without providing one for new process research. The categories were: new products, product improvements, new applications, and basic research. See *Chemical and Engineering News*, March 17, 1958, p. 52.

duce new products may find its demand and profits lower than its rivals', even if it should succeed in reducing the costs of producing its older products. This will be the case especially if the price elasticities of demand for its products are relatively low.

This explanation involves the question of profit maximization under conditions of oligopoly. Where profit rates are minimal, the pressures to reduce costs are associated with staying in business. Where profit rates are higher, however, the pressures to reduce costs via the introduction of new techniques are likely to be substantially less. But the pressures associated with the introduction of new products may be similar to those existing in the competitive model. Here the introduction of competitive new products by a firm's rivals may result in a large, and in some cases fatal, decline in the demand for the firm's output. While the conventional theory of profit maximization requires that the "carrot" be as effective as the "stick" in eliciting certain kinds of behaviour, it appears likely that the prevention of declines of profit is more important than the making of gains, that the maintenance of existing market shares through the introduction of new products is more important than the reduction of cost, and that a firm will work harder for the former purpose than for the latter. This means, merely, that in an uncertain world firms operate under some kind of minimax strategy. On this basis, firms will emphasize new products rather than new processes in their research efforts.

THE ROLE OF THE PATENT SYSTEM

United States law provides that patents on pharmaceuticals may be granted for new products as well as for new processes. Process patents, however, are a considerably weaker form of protection. As in most chemical industries, it is comparatively easy to modify a process somewhat and thereby evade the patent. Product patents, on the other hand, are a relatively strong vehicle of protection from competive suppliers. Once a product patent has been granted, a rival firm cannot supply the identical compound without fear of legal proceedings. As a result, nearly 80 per cent of medical patents are issued for new products rather for new processes.[23] Patent protection in this industry is, moreover, especially significant since a large proportion of individual products are covered, A survey of the industry estimated that over two-thirds of all prescription sales are for patented drugs.[24]

Although patent protection in the pharmaceutical industry takes the form primarily of product patents. it does tend to be limited to specific chemical substances. In most cases patents cover only a single compound or a small number of compounds. Especially when the advancement in knowledge is small, the scope of the patent may be rather limited. In addition, there is the question of patent specifications and the problem of anticipating all possible variants of the product.[25] With regard to pharmaceuticals, it is frequently possible to invent around existing patents; to find a variant which has not been specified, obtain a patent for it, and introduce it as a competing product.

Since a large proportion of pharmaceuticals have some degree of patent protection, entry into a specific therapeutic market requires, in most cases, some form of scientific or chemical product differentiation. In a world of competing monopolists, rivalry requires the ability to acquire a monoply position. The importance of pharmaceutical patents, however, can easily be overstated. While monopoly positions are conferred, patent protection does not normally confer the power to monopolize any of the therapeutic markets. This is borne out by the high turnover among leading firms and the vigorous product competition within these markets. The impact of the patent system has not been to create monopoly positions which remain active throughout the seventeen-year life of the patent, but rather it has been to foreclose to a great extent rivalry between identical chemi-

---

[23] In 1957 the industry's rate of return on investment after taxes enqualled 21.4 per cent, which placed it second on the F.T.C. list of thirty-nine industries. The comparable rate reported for All Manufacturing was 11.0 per cent. See Federal Trade Commission, *Report on Rates of Return for Identical Companies in Selected Manufacturing Industries, 1940, 1947–1957*, pp. 34–49, and Senate Report, *op. cit.*, pp. 53–55.

[24] In a sample of the twenty-two major pharmaceutical firms, for 1958 drug operations only, selling expenditures reached nearly 25 per cent of total sales. Research expenditures, on the other hand, equalled only slightly more than 6 per cent. Senate Report, *op. cit.*, p. 31.

[25] Arthur D. Little. Inc., *op. cit.*, pp. 11–30. While these ratios are probably over-stated, the extent of the over-statement is not likely to be by much more than one-fifth.

cal entities or standardized commodities about which price competition might develop. It has strengthened and encouraged the high degree of chemical product differentiation which is the primary form taken by technical change within the pharmaceutical industry.

THE ACHIEVEMENT OF MARKET POWER

The growth of effective product differentiation has led to an appreciable increase in industry profits and their maintenance at a relatively high level. In fact, the pharmaceutical industry has become one of the most profitable in the American economy.[26] The high level of market power, of which these profit rates are indicative, has evolved as part of the pattern of inter-relationship among the major firms, and is founded upon the achievement of product differentiation.

The crucial significance of product differentiation is that it provides the primary barrier to entry into the relevant therapeutic markets. Since effective entry normally requires some form of technical advance, the cost and risk of research comprise an important part of this barrier. Joined with research and development, moreover, are the extremely high selling expenditures undertaken by the larger firms.[27] Not only do these outlays accentuate the degree of differentiation among older products, but they also raise considerably the costs associated with launching a new product, and thereby provide a further barrier.

Entry barriers, created in this fashion, have resulted in fairly high levels of concentration within therapeutic markets. In a group of twenty such markets, the proportion of output accounted for by the leading five firms ranged from 56 per cent to 98 per cent.[28] It is within these markets that decisions on prices are made, and given such concentration ratios, we should not expect individual firms to disregard their own impact on market parameters. It is on this basis that market power has been achieved.

In this manner the leading five firms in the industry have maintained control of a large number of markets, and have created conditions within which the rivalry among themselves will proceed on a non-price basis. What is equally important, the achievement of product differentiation has succeeded in impeding the entry into these markets of the large number of smaller firms which would be likely to introduce some measure of price competition. Smaller firms have been forced to rely largely on standardized and non-patented products which are frequently non-competitive with the highly differentiated products which lead in most markets.

To consider further the role played by differentiation, it is instructive to examine the behaviour of the industry in patent licensing. Although pre-emption of the entire demand for a product seems most desirable for the firm, there are a number of factors which have increased the scope of licensing agreements. Not only may smaller firms wish to profit from the distribution facilities of their larger rivals,[29] but cross-licensing agreements may also be required to produce drug combinations and to settle "interference" proceedings.[30] In all of these cases, however, there appears to be a definite reluctance on the part of the major firms in the industry to license their smaller rivals even when licenses are granted to

---

[26] See the table compiled by the United States Patent Office of all patents relating to medicine issued during 1961. This table appears in United States Senate, *Drug Industry Antitrust Act*, Hearings before the Subcommittee on Antitrust and Monopoly, 87th Congress, 1st Session, 1961, Part 3, p. 1261. This document will be cited as *Antitrust Act Hearings*.

[27] *Ibid.*, Part 5, p. 2621.

[28] See the statement by George E. Frost, a member of the patent bar, in *Antitrust Act Hearings*, Part 4, p. 2119. Frost maintained with regard specifically to pharmaceutical patents that "the patent applicant is rarely able to anticipate and include all variants of his inventive concept in this document".

[29] The classic example is the case of meprobamate. The patent here is assigned to Carter Products, Inc., which markets the product under the name of Miltown. Within a month after the issue of the original patent, Carter licensed American Home Products Corp., one of the industry's largest firms with extensive selling and distribution facilities, to market the product under its own name, Equanil. Although sales of the latter quickly exceeded their own, Carter benefited through extensive royalties from the expansion of demand stimulated by the larger firm.

[30] An "interference" is declared by the Patent Office when a number of patent applications lay claim to the same invention. While the normal procedures in this case involve the Patent Office in an administrative hearing to determine the true inventor, the private settlement of claims is widely used in the pharmaceutical industry. This commonly results in the withdrawal of all but one of the original applications and the licensing of all parties when the patent is finally issued. See Senate Report, *op. cit.*, pp. 152–54.

other large firms. The motive is likely to have been the fear of instigating some form of price competition.

THE DIRECTION OF RESEARCH

The pharmaceutical industry during the post-war period has undertaken a heavy research effort which has been founded no a number of factors. The most important of these was an industrial structure characterized by both competitive and oligopolistic elements. The fact of vigorous price competition on the basis of standardized commodities led to a search for an effective method of achieving product differentiation. In addition, the development of a new technology, based originally on scientific advances made outside the industry, favoured the move towards a chemical form of differentiation. And finally, the existence of a strong patent system increased substantially the degree of differentiation which could be gained through the introduction of new products, and thereby strengthened considerably the incentive for research. These factors, however, have not only determined the extent of research in the pharmaceutical industry, but also they have influenced the type of work which is pursued and the framework within which it is undertaken.

Pharmaceutical research is conducted within a context of strong product rivalry. As a result a substantial amount of work is done which is designed primarily to invent around existing patents. While a number of important improvements have been made in this fashion, this work has led also to the development of a great many products which have pharmacological effects that are quite similar to those of drugs already on the market.[31] Moreover, a good deal of work has gone into developing new combinations of existing drugs and new dosage-forms. While research of this type is unlikely to result in major scientific advances, and frequently duplicates what has already been done, the minor advances and improvements which do arise are, nevertheless, important elements in the process of developing a new technology, and are essential if its full benefits are to be realized through the inevitable plethora of modifications and variations.

Product rivalry and the drive to strengthen product differentiation have led also in a second direction, which is concerned with the process by which new scientific knowledge is translated into marketable products. Competitive pressures have forced research directors in the industry to pay constant attention to new knowledge arising from external sources and have ensured that promising leads will rapidly be acted upon. This form of product development is vitally important because the major share of new scientific knowledge is likely to be produced in laboratories outside the industry. Not only do the research expenditures of the industry account for only about 30 per cent of total resources devoted to medical research within the United States,[32] but also considerable scientific activity is pursued abroad.

Even where the basic discoveries are made elsewhere, however, a good deal of work remains to be done.[33] The first problem concerns isolating or otherwise obtaining compounds which may have promising properties. It has been estimated, in fact, that approximately one-half of total pharmaceutical research and development expenditures is used to support the work of synthesizing, purifying, modifying, and preparing suitable substances for subsequent physiological tests.[34] Moreover, quite extensive biological and cilnical testing is required

---

[31] With regard to products in the latter group, the term "molecule manipulation" is commonly used because of the similarity in molecular structures which is frequently evident. An example of this feature of pharmaceutical research is provided by the case of three competing antibiotics: tetracycline, aureomycin, and terramycin. Although there may be some differences in their medical activity, these are not likely to be great, and the products may be considered as substitutes. Their molecular structures, although highly complex, are all rather similar. These structures are presented in Federal Trade Commission, *Economic Report on Antibiotics Manufacture*, 1958, p. 249.

[32] In 1960 the pharmaceutical industry allocated $215 million to research and development out of total national expenditures of $715 million. Government provided $400 million while other private support amounted to $100 million. These figures refer to the U.S. fiscal year 1960. *Antitrust Act Hearings*, Part 3, p. 1705.

[33] Despite the original discovery of penicillin by Alexander Fleming in 1928, thorough testing of the substance did not begin until 1939, and the development work, even under the exigencies of war, was not accomplished until 1943. It was only after the Second World War that penicillin became available on a significant scale to the civilian population. For a survey of the early discovery and development of antibiotics, see Federal Trade Commission, *op. cit.*, pp. 34–45.

[34] Statement by Austin Smith, President of the Pharmaceutical Manufacturers' Association, in *Administered Price Hearings*, Part 19, p. 10725.

before a new drug may be introduced. Research and development activities of this sort are extremely costly, although there is a reasonable degree of assurance of commercial viability. They demand extensive resources devoted to scientific efforts which, however, may be fairly routine. Thus, a major share of pharmaceutical research should be considered as complementary, rather than competitive, to research activities outside the industry.

TABLE 3.—NUMBER AND PERCENT DISTRIBUTION OF SCIENTIFIC AND PROFESSIONAL MANPOWER ENGAGED IN MEDICAL AND RELATED RESEARCH, 1960

A. DISTRIBUTION BY SECTOR

	Level of training			
	Total	M.D. and D.D.S.	Ph. D.	Other
Number	39,700	11,400	18,000	10,300
Percent:				
Government	19.7	17.3	12.6	34.5
Industry	18.1	4.8	12.2	43.2
Universities and research institutes	62.2	77.9	75.2	22.3
Total	100.0	100.0	100.0	100.0

B. DISTRIBUTION BY LEVEL OF TRAINING

	Number	Percent			
		Total	M.D. and D.D.S.	Ph. D.	Other
Government	7,800	100.0	25.3	29.2	45.5
Industry	7,200	100.0	7.6	30.6	61.8
Universities and research institutes	24,700	100.0	35.9	54.8	9.3
Total	39,700	100.0	28.7	45.3	26.0

Source: National Institutes of Health, "Manpower for Medical Research—Requirements and Resources, 1965–70," a report for the Committee on Appropriations, U.S. House of Representatives, February 1962, p. 24.

To consider this point further, it is useful to examine the data on educational levels of professional personnel engaged in medical research within and outside the pharmaceutical industry. The relevant information is presented in Table 3. As may be observed, 38 per cent of research professionals in the industry held doctoral degrees while 55 per cent of government researchers and 91 per cent of professionals within universities and research institutes had attained this level. Moreover, although the pharmaceutical industry provided approximately 30 per cent of total expenditures on medical research in 1960, it utilized only 5 per cent of total personnel with the M.D. degree and 12 percent of those with the Ph. D. degree. It appears, thus, that industry resources competed with non-industry resources for scarce, highly trained personnel to a lesser extent than would be indicated by the total size of the industry effort, and also that these resources were used primarily for alternative, non-competitive purposes.

The growth of competitive product differentiation has been associated with substantial outlays on research and development; but equally significantly, it has been associated with a specific emphasis in its direction. Research is a generic term which covers a broad spectrum of activities, and thus the type or character of the research activities undertaken as well as their volume are important factors in an analysis of the relationships among market structure, research and technical change.

[From The Review of Economics and Statistics, Vol. XLVII, No. 2, May 1965]

RESEARCH AND TECHNICAL CHANGE IN THE PHARMACEUTICAL INDUSTRY*

(William S. Comanor)

In recent years there has been a good deal of discussion concerning the relationships among market structure, research and development, and the rate of technical change. Much of this discussion has focused on the question of whether large firm size is a necessary condition before firms will engage in research, and whether research and development (R and D) is likely to grow more or less than in proportion to increases in firm size. A further set of questions deals with the relationship between research and the rate of technical change experienced by the firm. Can variation in the latter be explained largely by differences among firms in the size and character of their research programs? Are economies of scale in R and D likely to be present? What is the effect of firm size on the productivity of a research establishment? This paper provides an empirical analysis, concerned with these questions, of the experience of the United States pharmaceutical industry during the period between 1955 and 1960.

SOME MEASURES OF RESEARCH AND TECHNICAL CHANGE

The empirical investigation in this paper utilizes multiple regression techniques on a cross-sectional basis at the firm level. It examines the impact on technical change of a number of variables associated with the character of a firm's research and development effort. At this point, we will define the measures used to represent technical change and R and D, and describe the sample upon which the analysis is based.

Research and development establishments in the pharmaceutical industry [1] are concerned to a very large extent with the introduction of new products, and only meagerly with the development of new processes.[2] This is due to the fact that the primary motivating force behind R and D outlays in this industry lies specifically in the attempt to achieve scientific or chemical product differentation.[3] ucts at reasonable levels of cost, these are properly included under new product research, for we are concerned here not merely with the discover or invention of the product, but with the total R and D outlays, associated with its introduction. As a result, research output, in this paper, is measured entirely in terms of new products and we assume that the reesarch effort designed to produce new processes is small and can be safely ignored.

A concern with research output in terms of new products requires consideration of all products introduced which are new to the firm regardless of whether they have been introduced previously by competitors. This includes those products which are innovations or imitations and those which lie in an intermediate position, embodying varying degrees of differentiation relative to products already on the market. This last category, which in the pharmaceutical industry includes the great majority of new products, covers the entire spectrum of which pure innovations and pure imitations are the extreme values. Our measures of research output are, thus designed to represent the rate of new product technical change experienced by the firm.[4]

---

*This paper is based on the author's unpublished Ph.D. dissertation: "The Economics of Research and Development in the Pharmaceutical Industry," (Department of Economics, Harvard University, June 1963). I should like to express my appreciation to R. B. Heflebower, J. W. Markham, and L. D. Taylor for their counsel and advice during the course of my work on this project. I am also indebted to R. E. Caves, P. Dhrymes, W. E. Gustafson, J. Schmookler, and T. A. Wilson for a number of helpful suggestions, and to P. deHaen and R. A. Gosselin who generously provided the data used in this paper. The study was supported financially by the Science and Public Policy program at Harvard, and the computations were carried out under grants of subsidized time on I.B.M. 7090 computers at the Harvard Computation Center and at the M.I.T. Computation Center.

[1] In this study, we shall distinguish ethical drugs from proprietaries, and shall consider the pharmaceutical industry in terms of the former. Pharmaceuticals are marketed and sold only through the medical professions, and require, for the most part, a written medical prescription.

[2] So great is this emphasis that a 1958 trade journal survey of research budgets of pharmaceutical firms divided expenditures into four categories without providing a category for new process research. See *Chemical and Engineering News* (March 17, 1958), 52. Although funds are expended to develop techniques for mass producing new prod-

[3] See W. S. Comanor, *op. cit.*, 65–85.

[4] Technical change is considered to encompass imitation as well as innovation, and thereby deals with the entire process by which new technology is placed into actual practice.

New pharmaceutical products may be divided into a number of classes concerned with the innovative character of a new drug. In the first category are the single chemical entities which have not been introduced previously into United States markets. While these are largely innovative, they do include salts or derivatives of older products. Other classes deal with new combinations of active ingredients and new dosage forms of previously introduced products. These groups contain imitative as well as innovative characteristics. At the same time, many new products are purely imitative and merely duplicate products already on the market. In this study, we shall measure technical change by means of two variables. The first of these deals only with the introduction of new chemical entities, while the second includes a large imitative component and encompasses new products of all types.

New products vary not only in terms of the degree of innovative character which may be associated with them, but also in terms of their economic impact. To deal with this factor, we have weighed each new product on the basis of its sales during the first two calendar years following introduction.[5] We shall designate Y_1 to equal total sales, in their first two years, of all new chemical entities introduced by the firm during the period between 1955 and 1960, and Y_2 to equal a comparable sum when "new products" assume the broader definition.[6]

The scale of research and development has also been measured in two ways, both dealing with the number of persons employed in research establishments.[7] The first variable is the average number of professional R and D personnel employed in 1955 and 1960. The second deals with the average number of total personnel employed in R and D facilities at these two points in time. This latter value equals the sum of professioal and supporting personnel.

The sample used in this study includes 57 pharmaceutical firms,[8] and accounts for nearly 80% of total pharmaceutical prescription and hospital sales during the period between 1955 and 1960. Moreover, the sample is not dominated entirely by large companies, but rather covers the entire range of the size distribution of pharmaceutical firms.[9] The distribution of the sample is J-shaped with one-

[5] Sales data by product were obtained from a marketing research firm, R. A. Gosselin & Co., Inc., which supplied information for the years 1955 through 1961. These figures are derived from a representative sample of the nation's drug stores and hospitals. One shortcoming of these figures is that they do not represent total output but only that portion which is sold to consumers via prescriptions and to hospitals, and thereby do not include sales to government, corporations, or those made in physicians' offices. Prescription and hospital sales, however, account for approximately 80% of total pharmaceutical sales. See Lucy Kramer, "Drugs and Medicines" in *Public Health Reports* (Oct. 1958), 932. We are forced to assume that our conclusions would not be substantially altered if the remaining portion were included. Information on new products was obtained from Paul deHaen, a consultant to the pharmaceutical industry. In addition to providing information on year of introduction, these surveys noted the therapeutic classification of each new product and whether the product at introduction was a new chemical entity, duplicate, new compounded product, or new dosage form. For a more detailed discussion of the value and limitations of these data, see W. S. Comanor, *op. cit.*, 99–106.

[6] The data on pharmaceutical sales, used in the creation of these variables, have not been deflated for price changes but rather have been expressed in current dollars. This step was taken because of the high degree of price stability among pharmaceuticals during the period under consideration. See the price index computed by John M. Firestone in United States House of Representatives, *Drug Industry Antitrust Act*, Hearings before the Antitrust Subcommittee of the Committee on the Judiciary, 87th Congress, 2nd Session (1962), 608.

[7] Information on this subject for the years 1955 and 1960 has been published in National Academy of Sciences—National Research Council, *Industrial Research Laboratories in the United States*, tenth and eleventh editions. These surveys distinguish between professional and supporting staff within R and D establishments. In addition, for those companies in which the research effort was divided among a number of separate laboratories, the size and direction of research of each laboratory were, in most cases, listed individually. Since in many large diversified firms, the pharmaceutical R and D laboratories were distinguished from other research facilities, it was possible to limit our data for these firms to pharmaceutical research personnel. In the inevitable instances when we came across laboratories in which work was done on non-pharmaceutical as well as pharmaceutical projects, the facilities were included when it appeared from the summaries provided that the major part of the effort was concerned with pharmaceutical endeavors.

[8] The sample was chosen on the basis of data availability but with the condition that no substantial merger took place with other pharmaceutical firms between 1955 and 1960. Exceptions were made for those mergers which occurred during 1960 and in which the 1960 R and D survey reported separate and distinct research facilities. It was assumed, in these cases, that the pre-merger firms behaved and acted independently throughout the period.

[9] Size, in this study, is measured by the mean value of annual prescription and hospital sales between 1955 and 1960.

third of the observations falling in the smallest size class.

This sample is essentially non-random, and strictly speaking, our empirical conclusions apply only to it and not to any larger population. Since, however, the sample constitutes such a large proportion of industry output, and has a size distribution similar to one which might be expected from the entire industry, we may presume that conclusions based on an analysis of this group of firms probably can be applied to the entire industry.

THE MAJOR EMPIRICAL FINDINGS

In the analysis which follows, a quadratic term for research and development is introduced as well as a variable designed to represent the interaction between size of firm and scale of research establishment.[10] Some hypotheses concerned with the question of economies of scale in R and D distinguish between the productivities at a given level of research activity according to firm size. For this reason, a variable was computed which equated the product of scale of research and size of firm. Since R and D is measured on two bases, there are also two measures of the interaction between research and firm size. We may designate I_1 as the interaction variable where R and D is the number of professional, and I_2 where the number of total personnel is used. In addition, we include firm size as an explanatory variable and also a measure of output diversification.[11] The latter step enables us to test the relationship between this factor and the productivity of research. In the regression analysis, the measures of technical change and R and D are deflated by size of firm.[12]

As may be seen in table 1, the model fits the data far better when research is measured by professional rather than total R and D personnel. While we shall have more to say later about the question of supporting personnel, it appears that technical change is primarily associated with the number of professional investigators.

It may also be observed that diversification is negatively associated with our measures of technical change. When new products are defined in terms of new chemical entities, the coefficients are significant at the 99% level. With the broader definition of new products, however, the coefficients remain negative although there is some doubt as to their statistical significance.[13] It thus appears that diversification is more closely related to the introduction of new entities than to new products in general. In addition, the negative dimension of the parameters suggests that for given research and development, higher rates of technical change will be achieved if attention is concentrated towards a few product areas. To the extent that diversification of output denotes the scope of research activities, it may be that inefficiencies result from R and D undertakings which are "spread too thin" and that in the context of pharmaceutical research, it is better to work exhaustively with a limited number of problems.[14]

A further point is that Y_1 appears more closely associated with research and development than Y_2. We would expect this to be the case because Y_1 deals only with new chemical entities whose introduction is likely to have a relatively high degree of research input. Since Y_2 includes new chemical entities as well as other new products, the correlations observed with this variable may denote largely the influence of new entities.

---

[10] I am grateful to Lester D. Taylor for originally suggesting the use of an interaction variable.

[11] Diversification, in this analysis, deals only with the division of output among the various pharmaceutical markets and not with the division between pharmaceutical and non-pharmaceutical markets. On the basis of apparent medical usage, 40 therapeutic markets are defined. Our first measure of diversification, D_1, is the number of markets serviced by the firm which account for at least 2% of total pharmaceutical sales. The second measure, D_2, deals with the proportion of sales outside the firm's primary market, and equals one minus the ratio of sales in the firm's largest market to total pharmaceutical sales. The third measure, D_3, is the composite of the two previous ones. It is defined as the mean value of the products of D_1 and D_2, calculated on an annual basis. These measures are taken from Michael Gort. *Diversification and Integration in American Industry*, 8–11 and 23–24. In these regression equations, D_3 is the variable which is introduced.

[12] This step is taken to increase the likelihood that the assumption of homoscedasticity is satisfied.

[13] We should note that D_1 and D_2 were introduced into the analysis and provided quite similar results to those obtained from D_3. See W. S. Comanor, *op. cit.*, Table XXXI, p. 146.

[14] An alternative explanation of the negative sign of this coefficient concerns the possibility that there may be selling efficiencies which result from concentrating sales in a small number of therapeutic markets.

The relationship between research effort and technical change appears to be curvilinear and the quadratic term is always highly significant. The linear term, however, is not significantly different from zero and has an alternating sign when Y_2 is used to measure technical change, but is significantly negative when Y_1 is used. The existence of a negative cofficient for research and development effort in these cases does not result from the nature of the research process and the experience of the pharmaceutical industry, but rather is due to some statistical problems associated with the data concerned with the introduction of new chemical entities. These problems shall be discussed in the following section.

TABLE 1.—*Regression and correlation analysis—Summary findings*

A. RESEARCH AND DEVELOPMENT MEASURED BY PROFESSIONAL RESEARCH PERSONNEL

	Intercept	RD	RD²	Size	I_1	D	[a]R^2
(1) Y_1	[b] 0.422	[b] −4.671	[b] 0.547	[b] 0.0000344	[b] −0.000000128	[b] 0.130	[b] 0.40
	(0.136)	(1.285)	(0.107)	(0.0000083)	(0.000000031)	(0.040)	
(2) Y_2	[b] 0.873	−0.060	[b] 0.557	[c] 0.0000289	[c] −0.000000106	[c] −0.111	[c] 0.22
	(0.222)	(2.098)	(0.175)	(0.0000136)	(0.000000051)	(0.066)	

B. RESEARCH AND DEVELOPMENT MEASURED BY TOTAL RESEARCH PERSONNEL

	Intercept	RD	RD²	Size	I_2	D	R^2
(3) Y_1	[b] 0.471	[c] −1.989	[b] 0.112	[b] 0.0000300	[b] −0.000000051	[b] −0.120	[b] 0.28
	(0.150)	(0.981)	(0.031)	(0.0000102)	(0.000000017)	(0.044)	
(4) Y_2	[b] 0.932	1.381	[b] 0.118	0.0000194	−0.000000034	−0.100	0.18
	(0.232)	(1.513)	(0.047)	(0.0000157)	0.000000027)	(0.068)	

[a] Throughout this study, statistical significance for the regression coefficients is determined by one tailed t tests, and the significance of R^2 by the F ratio test.
[b] Indicates statistical significance at the 99-percent level.
[c] Indicates statistical significance at the 95 percent level.

THE ZERO VALUE PROBLEM

The data with which we are concerned cover a very wide range of variation in the scale of research facilities, and deal with the six years between 1955 and 1960 inclusive. Where new products include all introductions of whatever type, this period was sufficiently long to enable each firm to introduce at least one new product. In the case of new chemical entities, however, 17 firms, all with relatively small research establishments, failed to make even one introduction. As a result, the value taken by Y_1 for each of these 17 firms is zero. Although the zero values are clustered at the lower end of the distribution of research, when measured absolutely, they appear throughout the distribution when R and D is deflated by size of firm. If data for a longer time period had been available the number of zeros would most likely have been considerably reduced.

The existence of 17 zero values among the dependent variables raises a number of questions concerning the interpretation of the statistical results presented above. Although the usual least-squares properties apply to our estimates, some doubt is cast on the assumption that the disturbances are distributed normally, on which our tests of significance are based. Even more important, however, is the possibility of distortion resulting from the discontinuity which is introduced by a substantial cluster of values at zero.

The zero problem occurs only when Y_1 is used as the dependent variable. In no cases did Y_2 take the value of zero. Because of this fact, it is useful to compare regression coefficients between equations based on each variable. As may be noted in table 1, the coefficients in all cases but one are not widely different and are frequently quite similar. Only with regard to the first degree term for R and D do the coefficients diverge substantially.

To test for distortion in the equations concerned with Y_1, two techniques are used. The first is simply to eliminate the 17 zero values and compare the coefficients based on 40 observations with those based on 57. The second approach is concerned with examining regression equations estimated from grouped data.

Firms are aggregated in groups on the basis of research input, and an analysis is undertaken on group means.[15] In this manner, we are able to examine the relationship between research and technical change for firms even of the smallest size category without encountering the problems created by zero values in the dependent variable.[16]

Table 2 presents the results of a regression analysis in which the same relationship has been examined in four formats.[17] As may be observed, the equations in which the zero values are included and excluded present rather similar results. With the exception of the linear term for R and D, the regression coefficients do not take widely different values, although the standard errors of the coefficients are larger with the exclusion of the zero values. These parameters are highly significant. It does appear, however, that the significance and magnitude of the first degree R and D coefficient is peculiarly dependent upon the zero values.

The grouped regressions present greater difficulties of interpretation. Although the units were relatively homogeneous with regard to absolute values of R and D, there may have been considerable variation within group boundaries in terms of the explanatory variables used in the model. Grouping A, which has the effect of consolidating only the smallest research establishments, provides coefficients which are reasonable approximations of the ungrouped values. The estimates, with the exception again of the linear coefficient for R and D, remain highly significant; the intercept remains significant although not highly so. The second grouping format influences the relationship to a larger extent than the first. The coefficient of diversification becomes a good deal more negative and the standard errors of the parameters are increased substantially. As a result, only two estimates remain statistically significant.

With the exception of the first degree expression of research and development, it does not appear that our estimated regression coefficients have been distorted substantially by the inclusion of the zero values in the measure of technical change. It should be recognized, however, that we have not presented any theoretical explanation concerning the influence of these values.[18]

TABLE 2.—*Regression analysis—Examination of the zero problem*

	Intercept	R. & D.[a]	R. & D.[2]	Size	I_1	D	Number of observations
(1) Zero values included	[b] 0.422	[b] −4.671	[b] 0.547	[b] 0.0000344	[b] −0.000000128	[b] −0.130	57
	(0.136)	(1.285)	(0.107)	[b] 0.0000083	(0.000000031)	(0.040)	
(2) Zero values excluded	[b] 0.690	−13.584	[b] 0.545	[b] 0.0000306	[b] −0.000000118	[b] −0.163	40
	(0.227)	(10.284)	(0.144)	(0.0000114)	(0.000000041)	(0.055)	
(3) Grouping A	[c] 0.424	−3.604	[b] 0.627	[b] 0.0000422	[b] −0.000000152	[b] −0.169	29
	(0.219)	(3.522)	(0.153)	(0.0000128)	(0.000000045)	(0.066)	
(4) Grouping B	0.545	−2.249	[c] 0.538	[c] 0.0000430	−0.000000114	−0.237	16
	(0.695)	(6.625)	(0.268)	(0.0000225)	(0.000000088)	(0.194)	

[a] R. & D. is measured by the number of professional research personnel.
[b] Indicates statistical significance at the 99-percent level.
[c] Indicates statistical significance at the 95-percent level.

[15] Grouping was undertaken in two formats on the basis of non-deflated values of professional research personnel. The first format was to lay out the entire distribution of firms from smallest to largest in terms of the grouping variable and mark off boundaries so as to insure that at least 50 persons were included in each unit; grouping was begun at the small value end of the distribution. Twenty-nine groups were defined in this fashion of which all but ten consisted of single firms. In the second format, six groups were defined between the values of zero and 50, and intervals of 25 marked off throughout the remainder of the distribution. This scheme provided 16 groups.

[16] If we assume that the disturbances associated with individual firms have constant variances, then groups of varying size necessarily introduce heteroscedasticity. To retain the efficiency of our estimates, it has been necessary to weight all observations by the square root of the number of firms in the group; this also includes the column of ones used to determine the intercept.

[17] Since a grouping procedure in itself influences the degree of variation "explained" by a regression analysis on grouped observations, multiple correlation coefficients with grouping cannot be compared with those obtained from the original data. For this reason, the R^2's associated with these equations have not been presented.

[18] While a more sophisticated technique exists for dealing with this matter, the computational problems involved were too great to warrant its use. See James Tobin, "Estimation of Relationships for Limited Dependent Variables," in *Econometrica* (Jan., 1958), 24–36.

THE INFLUENCE ON TECHNICAL CHANGE OF SCALE OF RESEARCH AND SIZE OF FIRM

In the model presented above, an interaction variable is introduced to test whether firm size affects the shape as well as the position of the relationship between research and technical change. Since it is necessary to deflate by size of firm in order to deal with the problem of heteroscedasticity, this approach is used to reintroduce a scale factor and thereby examine differences in the form of the relationship as between small firms and large. In addition, an additive expression denoting firm size is used to consider the impact of this variable on the equation's position.

As may be observed in table 1, the linear expression of firm size is positively associated with research output. Since new products have been weighted by their sales during their first two years after introduction, we should expect our measures of technical change to be influenced by such factors as distribution facilities, selling effort, and firm reputation. To the extent that these factors are correlated with firm size, the size variable introduced into the equations will represent their influence, and the coefficient will be positive. Firm size appears to have acted to increase the gains resulting from new product introduction, and these gains increased more than proportionately with size of firm.

Our findings imply further that the relationship between research and development and the rate of technical change is substantially influenced by size of firm. Not only is the relationship shifted in total by differences in this variable, but also the functional form of the relationship is dependent on these values. Towards the lower end of the size distribution of firms, economies of scale in research seem to be present, while diseconomies are likely when firm size becomes large. It appears, thus, that we cannot deal with the question of economies of scale in R and D without considering the size of the firm within which the research is undertaken.[19]

These findings may be considered further by dealing with our primary relationship in a different format. If both sides of equation 1 in table 1 are multiplied by S, the relationship may be expressed in the form:

$$Y=a+bR+cR^2 \tag{1}$$

where,

$$b=-4.671-0.000000128\ S^2, \tag{2}$$

a is a function of S and D, and c equals 0.547.

Thus,

$$\frac{\partial Y}{\partial R}=b+2cR. \tag{3}$$

Since b is inversely related to S, $\frac{\partial Y}{\partial R}$ is larger for smaller values of S. For given values of R, $\frac{\partial Y}{\partial R}$ declines with the square of S. These results imply that for the pharmaceutical industry, the marginal productivity of professional research personnel is inversely related to size of firm.

Using the functional form of the relationship presented in (1), the elasticity of Y_1 with respect to R was calculated to three points in the distribution of S.[20]

[19] One additional point concerning our measures of technical change should be mentioned. Since these variables have been deflated by size of firm, they may be taken to represent the proportion of sales accounted for by new products. If larger firms sell relatively more standardized or older products than smaller firms, then we should expect the deflated ratios denoting technical change to be inversely related to firm size, and this fact would have little to do with the productivity of research establishments. It appears, however, that new products account for as large a proportion of total sales in large firms as in small ones. The proportion of total sales accounted for by products introduced within the preceding five years was computed for each year between 1955 and 1960 for each of the 57 firms included in the sample. The simple correlation coefficient (r) between the mean proportion for each firm and the corresponding measure of firm size was +0.18. The simple correlation coefficients between deflated values of Y_1 and Y_2 and firm size were also computed. The coefficients were +0.14 and +0.08 respectively. Since it does not appear that larger firms sell, on the average, proportionately less new products than smaller firms, we should not expect this consideration to distort our findings.

[20] The elasticity was calculated as $\frac{\partial Y}{\partial R}\cdot\frac{R}{Y}$.

Since this elasticity is defined as the percentage change in research output for a unit percentage change in R and D, holding S constant, it will describe the findings on economies of scale in research which are implicit in the equation results. Elasticities were computed at values of S equal to 1,000, 10,000, and 50,000 (pharmaceutical sales in thousands of dollars).[21] When S equaled 1,000, the calculated elasticity was 1.39. At S equal to 10,000 and 50,000, the corresponding figures were 0.61 and 0.54. While there are likely to be increasing returns to scale in R and D at low values of S, decreasing returns seem to be the case when S becomes moderately large.[22] [23]

SOME ALTERNATIVE HYPOTHESES

Up to this point we have proceeded on the assumption that the direction of causation runs from research and development to technical change, and have implied that the degree of diversification affects the rate of technical change. It is necessary, however, to consider some alternative hypotheses. It may be that successful new product introduction influences the extent of research activities rather than or in addition to the opposite alternative. A rapid rate of technical change may stimulate the firm to expand the scale of R and D facilities; accelerated technical change may result in higher profit rates which enable the firm to increase its investment in research and development. Similarly, it may be hypothesized that the rate of new product introduction is the vehicle by which diversification proceeds. If these alternative hypotheses are correct, we should expect that technical change at the start of the period would be better correlated with R and D and diversification at the end, than if the lead-lag relationship were reversed. Our original hypotheses would postulate that R and D and diversification should lead the variables denoting technical change. To investigates these questions, we shall consider our regression equations with alternative lead-lag structures.

The empirical findings concerned with this matter are presented in table 3. The research and development variables denote the number of professional research personnel employed in either 1955 or 1960, while the technical change and diversification variables span the period which is designated. The size variable remains the same as in the previous analyses.

The results are striking. With a lead-lag structure consistent with our original hypotheses, all coefficients are significant at the 99% level. With the opposite lag structure, the R^2's are not significant, the parameters of diversification are smaller than their standard errors and alternate in sign, and the coefficients for size, interaction, and the quadratic expression for research are not significant. There is, however, some positive linear association between technical change in 1955–1957 and research in 1960 when the variable Y_2 is used, but this relationship may be relatively weak for it fails to appear in the regression equations dealing with the entire six-year period. We may conclude, therefore, that the apparent lead-lag structure is not inconsistent with our original hypotheses.

---

[21] Corresponding values of R were determined by taking the mean value of that variable for firms in the neighborhood of the appropriate value of S. For S equal to 1,000, firms were included between S equal to 500 and 1,500; for S equal to 10,000, the corresponding figures were 5,000 and 15,000; while for 50,000, the firms with S between 40,000 and 60,000 were included. The three values of R which were determined in this fashion were 13.1, 59.2, and 353.3, respectively. D was computed at its mean value throughout.

[22] One qualification of the above results should be mentioned. It may be that large firms undertake more "basic research" than smaller firms. Although it is assumed that the returns from this activity are obtained through its advancement of the rate of new product technical change, the time horizon in this case may be substantially longer than that of "applied research and development" so that the six-year period under examination may not be long enough to measure accurately the returns. As a result, the level of research output for larger firms would be understated relative to smaller firms. It does appear that larger firms spend somewhat larger proportions of their R. & D. budgets for "basic research." For 1959, National Science Foundation data for industry group "Drugs and Medicines" (SIC 283) disclose that firms with total employment over 5,000 undertook 50% of total industry funds for the performance of R. & D. while these firms accounted for nearly 63% of industry "basic research." The difference between the elasticities is, therefore, likely to be somewhat overstated.

[23] One further implication of the functional form of (1) should be mentioned. Since b is negative and c is positive, the function, at given values of S, originally declines, reaches a minimum, and then increases. This minimum is reached at values of R equal to $\frac{-b}{2c}$. In addition; since b is an inverse function of S, the values of R at which the minimum is reached increases with S. At values of S equal to 1,000, 10,000, and 50,000 the minimum is reached at values of R equal to 4.4, 16.0 and 298. Actual values of R will, in nearly all cases, be greater than these values. At the extreme upper tail of the distribution of S, however, the marginal productivity of research estimated by equation (1) in table 1 will be negative.

TABLE 3.—*Regression and correlation analysis—Various lead-lag structures*

	Intercept	RD[a]	RD[2]	Size	I_1	D	R^2
		1955	1955		1955	1955–1957	
(1) Y_1 (1958–1960)	[b] 0.357	[b] −3.360	[b] 0.351	0.0000159	[b] −0.0000000693	[b] −0.109	[b] 0.31
	(0.100)	(1.025)	(0.083)	(0.0000049)	(0.0000000212)	(0.032)	
		1960	1960		1960	1958–1960	
(2) Y_1 (1955–1957)	0.125	−1.193	0.0861	0.00000649	−0.0000000172	−0.0164	0.13
	(0.076)	(0.751)	(0.0519)	(0.00000485)	(0.0000000152)	(0.0205)	
		1955	1955		1955	1955–1957	
(3) Y_2 (1958–1960)	[b] 0.594	[b] −3.898	[b] 0.465	[b] 0.0000171	[b] −0.0000000810	[b] −0.118	[b] 0.32
	(0.122)	(1.252)	(0.102)	(0.0000060)	(0.0000000259)	(0.039)	
		1960	1960		1960	1958–1960	
(4) Y_2 (1955–1957)	[c] 0.311	[b] 4.143	0.0053	0.00000309	0.0000000012	0.0131	0.18
	(0.140)	(1.387)	(0.0959)	(0.00000895)	(0.0000000280)	(0.0378)	

[a] R and D is measured by the number of professional research personnel.
[b] Indicates statistical significance at the 99 percent level.
[c] Indicates statistical significance at the 95 percent level.

THE ROLE OF SUPPORTING PERSONNEL

At this point we shall examine the question of whether a large ratio of supporting personnel to professional staff substantially increases the efficiency of a research facility. The relevant variable is defined as the mean ratio in 1955 and 1960, and the analysis is carried out by introducing it into our regression equations.

The coefficients of the ratio of supporting to professional personnel are always negative and in no instances are they statistically significant.[24] Thus, we may conclude that supporting personnel play only a minor role in determining the magnitude of research output. High ratios of auxiliary staff to professionals do not seem to increase the productivity of the professional researchers. These results corroborate our earlier observations that the number of professionals seems to be a better measure of research input than the total number of R and D personnel.

THE EFFECT OF GROWTH IN RESEARCH FACILITIES

In this section, we examine the influence on technical change of rapid expansion of R and D activities. Some observers maintain that research efficiency is likely to be impaired if the rate of expansion is large.[25] To consider this position, we define two variables and introduce them separately into the regression equations. The first variable, which is designated G_1, is the ratio of total research personnel in 1960 to the corresponding value in 1955. The second variable G_2, is the ratio of 1960 to 1955 levels of professional research personnel.

The coefficients of G_1 and G_2 are not statistically significant although they are negative and thereby do take the hypothesized sign.[26] As a result, the hypothesis was restated and examined in a different format. It may be that rapid expansion of research facilities does not affect the rate of technical change continuously. Inefficiences may be created only when growth is pushed above some threshold rate. To investigate this position, three dummy variables are defined according to whether the research establishment showed no growth, moderate growth, or rapid growth, between 1955 and 1960, and the first two dummies are introduced into the regression equations.[27]

[24] When the ratio of technicians to professionals is introduced into equations 1 through 4 of table 1, the estimated coefficients and standard errors are: (1) −0.181, 0.144; (2) −0.149, 0.220; (3) −0.086, 0.139; and (4) −0.161, 0.211. (In equations 2 and 4, the non-significant linear RD variable was omitted when these estimates were obtained.)

[25] Mansfield has reported: "There are considerable costs involved in a very rapid expansion of a firm's R and D department, the importance of which was stressed in interviews with various executives." Edwin Mansfield, *The Expenditures of the Firm on Research and Development* (mimeographed), 4.

[26] When G_1 is introduced into our original regression equations, the estimated coefficients and standard errors are: −0.124, 0.126 in the case where technical change is measured by Y_1, and −0.168, 0.185 when Y_2 is used. The corresponding estimates for G_2 are: −0.0229, 0.0958 and −0.061, 0.153. In these calculations, RD is measured by the number of professional research personnel.

[27] The dummy variables are defined on the following basis: If, during the period, the laboratory remained stable or declined in size, the firm is listed in the first category; if the facility increased in size but by less than 100%, the moderate growth grouping is designated; while if the research establishment at least doubled in size, it is considered to have experienced rapid growth. The three categories are defined in terms of increases in both total and professional research personnel.

An F test is used to deal with the question of whether the dummy variables significantly affect the rate of technical change. While the dummies based on increases in professional personnel do not lead to significant values of F even at the 50% level, those based on the expansion of total research personnel result in test values which are significant at the 75% level when Y_1 is used to measure technical change. In addition, the estimated regression coefficients founded on the latter set of dummies take the expected sign and relative order of magnitude.[28]

These findings point out that any inefficiencies resulting from rapid expansion of research facilities are more likely to be associated with increases in the total size of the establishment rather than particularly related to increased hiring of professional personnel. It should be emphasized, however, that empirical support for this hypothesis exists only if we are willing to accept statistical findings as the 75% level of significance. As a result, this conclusion is highly tentative.

CONCLUDING STATEMENTS

In this paper, we have investigated the relationship between research and development and the rate of new product technical change in the pharmaceutical industry. From the empirical evidence, there appears to be a fairly sustained association between research input and new product output. Within this industry, research expenditures are not undertaken merely with the hope of some distant but unknown returns, but rather with the expectation that profitable gains will accrue within a reasonable period of time.

Our analysis also provides some evidence that in the pharmaceutical industry, there are substantial diseconomies of scale in R and D which are associated with large firm size; and that these disadvantages are encountered even by moderately sized firms. One implication of this finding is that an actively enforced pro-competitive policy in this sector is not likely to dampen the rate of technical change and may well stimulate it. While little is known about the extent to which this result is applicable to the economy at large, it does appear that there are grounds for considerable doubt as to the position that large firm size is always a necessary condition for rapid technical advance.

[From The Journal of Business of the University of Chicago, Vol. XXXIX, No. 1, Part I, January 1966]

THE DRUG INDUSTRY AND MEDICAL RESEARCH

THE ECONOMICS OF THE KEFAUVER COMMITTEE INVESTIGATIONS*

(By William S. Comanor†)

In the course of recent American politics, few confrontations have been more lively and more prolonged than that which was waged between the pharmaceutical industry and the Senate Subcommittee on Antitrust and Monopoly (The Kefauver Committee).[1] In December, 1959, the committee opened hearings on the

---

[28] The dummy variables denoting "no growth" and "moderate growth" are introduced into equations (1) and (2) of table 1. When the dummies describe the rate of expansion in total research personnel, the estimated coefficients and standard errors are, for equation (1): 0.371, 0.220 and 0.209, 0.184, and for equation (2): 0.399, 0.336 and 0.159, 0.297. The coefficient of the "no growth" dummy in equation (1) is statistically significant at the 95% level. Furthermore, the intercept, which encompasses the effect of the third classification, is reduced from 0.422 to 0.199 by the introduction of the dummies into equation (1).

*This article is drawn in part from my unpublished Ph.D. dissertation, "The Economics of Research and Development in the Pharmaceutical Industry" (Department of Economics, Harvard University, June 1963). I am grateful to J. W. Markham and R. B. Heflebower, who acted as supervisors of the original study, and to Carl Kaysen for valuable counsel in the writing of this paper. I have also benefited from comments by Oswald H. Ganley, Mary Lee Ingbar, Lars Sandberg, and Harvey M. Sapolsky, and from the secretarial assistance of Anne Caneles. This study was undertaken with the support of the Science and Public Policy Program at Harvard.

†Instructor of economics, Harvard University. Currently on leave and associated with the Antitrust Division, U.S. Department of Justice. This article was written at Harvard and in no way does it reflect the views of the Department of Justice.

[1] The pharmaceutical industry in the Congressional proceedings, and therefore in this paper, is limited to firms that produce ethical drugs, as opposed to proprietaries, and that distribute these products in dosage form. Pharmaceuticals, thus, are marketed and sold through the medical profession and require, for the most part, a written medical prescription.

industry, and 11,000 pages later, testimony was still being recorded. Now that the record has been closed and the dust from the contest has been allowed time to settle, it is useful to consider some of the questions raised in the proceedings. This paper reviews the evidence on drug research.

I

A major point at issue between the industry and the committee concerned the medical value of industry research. Although forced to admit that expenditures on research were large,[2] the committee charged that little of social value came from industry laboratories. It maintained that nearly all the important new discoveries in recent years had been derived from research work performed outside the industry and that commercial laboratories are concerned primarily with "molecule manipulations" or new drugs therapeutically quite similar to drugs already on the market.[3] The committee implied that industry research was highly duplicate and that a large portion of this work could be eliminated without reducing very much the flow of *important* new drugs. In defense of its efforts, the industry argued that from industry laboratories come most of the new drugs that are extensively utilized in current medical practice.[4]

These opposing views on the nature and significance of the pharmaceutical research effort collided in what may be termed the "battle of the lists." This contest was based on competing lists of new drugs prepared by the committee and the industry. One of the major points of difference between the two is that the committee tended to concentrate on drugs that embodied what it considered to be major therapeutic advancements at the time of introduction. As we would expect, the committee and its lists emphasized the role played by foreign and non-industry laboratories.[5] The industry list, on the other hand, included new drugs that may not have embodied large steps forward but that are in frequent use and thereby seem to have the confidence of the country's physicians.[6] A large majority of the drugs on this list were discovered and developed within industry laboratories. While there was some controversy over the origins of specific drugs, it seems likely that both sets of lists and their associated conclusions are essentially correct.

To unravel the issues here, it is necessary to deal explicitly with the nature and functions of industry research. In the succeeding sections of this paper, we shall undertake this discussion in the light of the alternative positions taken by the committee and the industry.

II

Research and development in the pharmaceutical industry is carried on within a context of strong product rivalry, and this factor has affected the type of work that is undertaken. Considerable attention is focused on the rapid introduction

[2] In 1963, e.g., company-financed research and development totaled $282 million. This sum equaled nearly 9 per cent of total industry sales (Pharmaceutical Manufacturers Association, "Pharmaceutical Industry Research and Development Activity, 1963–64" [Washington, D.C., 1964 (mimeographed)], p. 1, and Pharmaceutical Manufacturers Association, "Ethical Pharmaceutical Sales, 1963" [Washington, D.C., 1964 (mimeographed)], p. 2). Although N.S.F. data on company-financed R & D as a percent of net sales in manufacturing industries for 1963 is not yet available, 1961 data show that the highest percentage for an industry other than drugs and medicines was 4.4 per cent (National Science Foundation, *Research and Development in Industry*, 1961 [February, 1964], p. 86).

[3] This position was based on statements such as the following, which appeared throughout the proceedings: "The question is what then is the goal of this admittedly large-scale laboratory effort of our industry? Partly to exploit and market those foreign and non-industrial advances and compounds that I have mentioned. Mostly, however, to modify the original drugs—just enough to get a patentable derivative, but not to change it enough to lose the original effect" (Testimony of Frederick H. Meyers, associate professor of pharmacology, University of California, in U.S. Senate, *Administered Prices in the Drug Industry* [hereinafter cited as "Administered Price Hearings"]. [Hearings before the Subcommittee on Antitrust and Monopoly (Washington: 86th Congress, 2d Session, 1960)] (Part XVIII, p. 10394).

[4] Pharmaceutical Manufacturers Association, *The Pharmaceutical Industry, Key Facts at a Glance* (Washington, D.C., n.d.). The industry reported that of the ninety-four important drugs developed since 1945 and in frequent use in 1960, forty-nine, or over half, were discovered by drug companies or by university scientists financed by drug companies.

[5] The original committee list appears in *Administered Price Hearings*, Part XIX, pp. 10943–45, while the supporting comment is made in U.S. Senate Subcommittee on Antitrust and Monopoly, *Report of the Study of Administered Prices in the Drug Industry* (Washington: 87th Congress, 1st Session, 1961), p. 115.

[6] E.g., the original industry list included Achromycin as well as Terramycin, two antibiotics that have similar therapeutic effects. This list appears in *Administered Price Hearings*, Part XIX, pp. 10840–54.

of new products, and, as a result, the number of introductions exceeded 3,800 in the decade ending in 1960.[7] Of these, the largest number by far consisted of new combinations of existing drugs and of new dosage forms.[8] The remainder, or 432 in total, were new chemical entities, and these absorbed the largest share of the research effort. It should be noted, however, that this category included salts and other derivatives of known drugs as well as entirely new substances.[9]

The nature of industry research may be inferred to some extent from information on new products. As we might expect, there appears to be a large effort to invent around existing patents, and this has led to the introduction of a considerable number of new drugs whose therapeutic effect is quite similar to that of products already on the market. We should note, however, that rarely do different chemical compounds have perfectly identical therapeutic properties and that small differences may be important for some patients.

To a great extent, industry research activities build upon the scientific achievements of the past. They are concerned, primarily, with the smaller, albeit less uncertain, steps forward. Although these activities may duplicate what has already been done and also frequently may lead to new drugs that represent at best a very minor advance, it is still true that a number of important modifications and variations have been discovered within industry laboratories.

The second major emphasis of industry research concerns the process by which new scientific knowledge is translated into marketable products. Product rivalry is strong, and this has insured that new drugs will not lag far behind advances in scientific knowledge, no matter whether these advances originate in university, government, or foreign laboratories. This is an important area, because in most cases substantial research and development problems remain even after the original discovery has been made. Not only is there frequently a substantial gap between laboratory synthesis and large-scale manufacturing methods, but also considerable analysis and testing is required before a drug may be marketed. And these activities, which may well be more costly than basic research, are undertaken for the most part within industry laboratories.

Whether the impetus for a new drug comes from competitive successes, from basic research carried on within the industry, or from new scientific knowledge arising from non-industry sources, the problems of synthesis and testing remain, and these absorb the major share of industry research expenditiures. Large quantities of the promising compound are normally required before efficient production techniques are available, and large sums are expended for this purpose. It has been estimated, in fact, that approximately one-half of total expenditures on research with the industry goes to support the work of synthesizing, purifying, modifying, and preparing suitable substances for subsequent physiological tests.[10] Once this process is completed, the new substance is subjected to intensive biological investigation. The first concern is to determine the therapeutic properties of the substance this encompasses the major purpose of the drug as well as side effects. In addition, it is necessary to test for toxicity levels and to determine the compound's potency in order to gain knowledge of appropriate dosages. For these purposes, the pharmaceutical industry used in its research facilities some nine million laboratory animals during 1961.[11]

There is, moreover, a great deal of uncertainty concerning the therapeutic properties of new substances. As would be expected, most of those tested in laboratories are found to be without sufficient promise to justify clinical testing on human beings. In 1958, for example, nearly 115,000 substances were subjected to biological tests by pharmaceutical firms, while only 1,900 were considered worth testing clinically.[12] The Food and Drug Administration requires that manufacturers undertake efficacy studies on new drugs and that the results of these tests

---

[7] Paul de Haen, "New Product Survey" (New York: privately printed by author, 1960).

[8] Although new chemical substances are not involved in these cases, considerable attention may still be required. E.g., a good deal of work was done in an attempt to combine Glucosamine with certain antibiotics in the hope that this would increase the absorption of the drug into the blood stream (*Administered Price Hearings*, Part XVIII, p. 10257).

[9] During 1959, sixty-three new chemical entities were introduced. Of these, eleven were new salts of old products and twenty-three were derivatives of known drugs. The remainder of twenty-nine were entirely original products (U.S. Senate, *Drug Industry Antitrust Act* [hereinafter cited as "Antitrust Act Hearings"] [Hearings before the Subcommittee on Antitrust and Monopoly (87th Congress, 1st Session, 1961)], Part II, p. 888).

[10] Statement by Austin Smith, president of the Pharmaceutical Manufacturers Association, in *Administered Price Hearings*, Part XIX, p. 10725.

[11] Reported in *New York Times*, August 26, 1962, p. 60.

[12] *Administered Price Hearings*, Part XIX, p. 10725.

be submitted to the administration before a new chemical entity may be introduced.

Clinical studies are normally carried out with the help of physicians and medical institutions throughout the country. New drugs are contracted out to clinicians in medical schools or hospitals for controlled testing. While there may be a sizeable liaison group among the pharmaceutical research staff, the actual testing is done, in most cases, by non-industry personnel. In 1955, only one major company supported directly the maintenance of a clinical-research group having hospital facilities; this was the Lilly Laboratory for Clinical Research located within the Indianapolis City Hospital.[13] Even when the testing was performed outside of industry facilities, however, the pharmaceutical firm concerned normally supplied the compound that was to be tested.

While it is difficult to determine directly the cost of clinical trials to pharmaceutical firms, we can observe the division of industry research expenditures spent within and outside of the companies as some indication of the relevant magnitude. In this regard, less than 10 per cent of total R & D expenditures was in the latter category.[14] To this should be added, however, the cost of preparing the compounds used in the trials.

Although the laboratory and clinical-research activities described above have a highly applied and developmental character, still they are necessary before medical benefits can be realized from new scientific knowledge. With regard to the major share of industry-research activities, these are best viewed as a complement rather than as an alternative to those undertaken outside the industry.

III

To evaluate industry research activities, it is necessary to place them within the context of the total medical research effort. In this regard it should be stressed that the industry provides only about 30 per cent of total medical-research outlays within the United States,[15] and the proportion would fall still further if foreign research expenditures were included in the total. Given the relative size and essential complementarity of industry and non-industry research, it follows that the two cannot adequately be evaluated by the same set of standards. Despite the fact that the same word is used to describe both sets of activities, they are essentially different.

One of the major problems of the investigatory proceedings and the joust between the committee and the industry concerned the failure of either side to recognize this difference. The committee criticized the industry on the apparent failure of industry scientists to provide many dramatic new breakthroughs on which major therapeutic advances are based and also on the lack of important new drugs in recent years. Whatever the validity of these charges, they are not founded on a realistic appraisal of the role and function of industry research. At the same time, however, the industry's apparent claim that its research is responsible in large measure for the great advances in medical science that have been achieved in recent years is equally inappropriate. Both of these positions imply that industry and non-industry research can be evaluated on a single set of standards, and this neglects the considerable differences in purpose and approach between the two areas of work.

A valid judgment of industry research and development must necessarily be founded upon the nature of the complementarities present within the larger boundaries of medical and health-related research. These complementarities are important, and it does appear that the differences that exist between the two areas of research may well constitute a necessary and desirable division of labor. The peculiar attributes of each probably create for it a comparative advantage for the specific activities which it, in fact, undertakes. While it is difficult to be definite in these matters, it does seem likely that basic or fundamental research is better carried on in a university or government laboratory

[13] American Foundation, *Medical Research, A Mid-Century Survey* (2 vols.; Boston: Little, Brown & Co., 1955), I, p. 592.

[14] In 1958 the amount spent outside of industry facilities was $13.4 million out of a total of $170.0 million, while in 1959 the comparable figures were $15.3 million and $190.0 million (*Administered Price Hearings*, Part XIX, p. 10724).

[15] In fiscal year 1960, the pharmaceutical industry allocated $215 million to research and development activities out of a total national expenditure of $715 million. Government provided $400 million, while other private support equaled $100 million (*Antitrust Act Hearings*, Part III, p. 1705).

than in a commercial facility. Within a profit-motivated firm, there is a latent conflict between an atmosphere of freedom, which is necessary for good scientific endeavor, and the necessity of direction from those who are responsible for the conduct of the firm.[16] In addition, it appears that those firms which have had the greatest success in basic research have moved furthest toward creating a university environment within their facilities.[17] While additional corroborating factors could be mentioned, it is sufficient to remark that we should not be surprised if the major advancements in medical science are achieved in laboratories outside the pharmaceutical industry.

At the same time, moreover, there are a number of reasons why highly applied research and development activities are best carried on within the confines of the industry. Not only do these tasks assume a highly routine character which tends, in a large number of cases, to make them uninteresting to university scientists, but also the relevant procedures are sufficiently well understood so that they may be reasonably well directed. An atmosphere of completely free inquiry is not crucial where both the goals and the procedures are relatively well defined. These considerations, moreover, may explain the fact that although the pharmaceutical industry provided approximately 30 per cent of the total funds for medical research in 1960, it utilized only 12 percent of the total number of persons doing medical research with the Ph.D. degree and only 5 per cent of the total number doing medical research with the M.D. degree.[18]

Even more important, however, are the economies that are likely to result from combining applied research with production in the same organization. Within a technical industry, the road from laboratory to factory is not level or direct, and large gains may result from institutional arrangements that provide the smallest possible barriers to the flow of information. To insure a free and unrestricted flow it is probably necessary that the latter stages of research and development, as well as production, are undertaken within the firm.

It is interesting, in this regard, to note the conclusions of a recent study that contrasted the American and Soviet pharmaceutical industries. Unlike the United States, research and production are carried on by completely separate agencies in the Soviet Union. The authors report that as a result of this arrangement, "there are difficulties of communication and coordination between [research] institute and factory which constitute a major bottleneck in getting pharmaceuticals into production."[19] While the divorcement of research from production represents merely one additional factor out of many that distinguish the Soviet industrial structure from that of the United States, still the apparent results from this divorcement do seem to point to a major set of problems.

The division of labor that exists between industry emphasis on product development and non-industry emphasis on the more fundamental areas of research provides a system that is rational in principle although we know little of whether the optimal share is more or less than the 30 per cent presently occupied by the industry. If, however, our concern is with the use and possible misuse of such scarce and limiting factors as highly trained personnel rather than merely with dollars, the quantitative problems appear in a different perspective. A large drug-industry effort does not divert, to a substantial degree, scarce resources from university and government laboratories to its own facilities but, rather, uses, for the most part, a different and non-substitutable class of personnel. Thus, the opportunity cost of industry research in terms of non-industry research is likely to be relatively low.

When evaluated on its own terms, it seems clear that the industry accomplishes an important research task in a generally effective manner. It has been responsible for a large number of pharmacological modifications and improvements that have been introduced as well as for an acceleration in the process of testing and developing new drugs, and both of these gains provide a high degree

---

[16] This "conflict" is discussed in John Jewkes, David Sawers, and Richard Stillerman, *The Sources of Invention* (New York: St. Martin's Press, 1958), chaps. vi, vii.

[17] See, e.g., the statement by James B. Fisk, "Basic Research in Industrial Laboratories," in Dael Wolfle (ed.), *Symposium on Basic Research* (Washington, D.C.: American Association for the Advancement of Science, 1959), pp. 159–67.

[18] National Institutes of Health, *Manpower for Medical Research—Requirements and Resources, 1965–1970.* (A Report for the Committee on Appropriations [Washington: U.S. House of Representatives, February, 1962]), p. 24.

[19] Raymond A. Bauer and Mark G. Field, "Ironic Contrast: U.S. and U.S.S.R. Drug Industries," *Harvard Business Review* (September–October, 1962), p. 93.

of social value.[20] It should be recognized, however, that these gains are on a very different scale from those related directly to significant advancements in medical science and to dramatic improvements in public health.

IV

During the investigations, the Kefauver Committee proposed a major revision in economic policy toward the drug industry. This proposal, however, was rejected by the parent Committee on the Judiciary and was not included in the 1962 Drug Act. The change would have instituted compulsory licensing at reasonable royalty rates three years after a patent had been granted. It is useful here to note the implications of this change in the light of the analysis presented above.

Although this proposal was intended to produce more competitive levels of prices and profits, the industry maintained that it would also lead to much reduced research expenditures. This position seems to be generally correct. The primary motive for large research efforts in this industry has been the drive to achieve effective product differentiation.[21] With compulsory licensing, there would be a sharp decline in the extent of differentiation based currently on chemical differences among products.

Nevertheless, it is not at all certain that compulsory licensing would significantly lower the rate of introduction of the most important new products. The largest proportion of these come originally from non-industry laboratories. It is true that pharmaceutical companies would have less incentive to undertake projects of a long-term nature, such as many of those in basic research, because of the diminished prospect for large gains over a prolonged period of time. These firms account, however, for only a minor share of the work which currently is done, and they also appear to lack a comparative advantage in pursuing basic research in an extensive manner.

Even with compulsory licensing, research and development would still comprise an important element of industry behavior. There would still be gains from achieving product differentiation. New drugs that embody a large element of therapeutic improvement are also likely to provide a high degree of product differentiation. Having developed a differentiated product, not only would firms benefit substantially from a head start in promotion and selling, but also this advantage would last for the initial, prelicense period of patent protection, during which time monopoly gains could still be attained. In addition, patent royalties might become an important element of the rewards resulting from successful research.

Compulsory licensing appears on balance to be a useful and desirable policy to adopt. While the magnitude of research would decline, it does seem probable that projects of limited medical value and of lowest industry comparative advantage would be eliminated first. There are substantial social gains to be derived from industry research, but the marginal social productivity of research may well decline rapidly after a certain level has been reached. We should be wary of believing that much is to be gained from ever higher levels of research and development, and it is quite possible that the present effort may exceed that required to fulfill the major research functions and responsibilities of the industry. So long as the decline in research expenditures was not of overwhelming proportions, it may well be a small "price" to pay for a more competitive determination of pharmaceutical prices and profits.[22]

---

[20] This is in contrast to the Soviet pharmaceutical industry. Bauer and Field state that "the testimony of the well-informed Soviet sources is that the separation of research from production tends to produce substantial delays in the availability of drugs to physician and patient" (Bauer and Field, *op. cit.*, p. 94). At the same time, however, excessively rapid process leads to new drugs that have not been adequately tested, and this charge has been laid on the doorsteps of the American pharmaceutical industry. This point was commented on by a medical educator who noted sharply: "There is no short cut from chemical laboratory to clinic, except one that passes too close to the morgue" (*Administered Price Hearings*, Part XVIII, p. 10418).

[21] The role and function of product differentiation and its relation to research and development activities is discussed in William S. Comanor, "Research and Competitive Product Differentiation in the Pharmaceutical Industry in the United States," *Economica* (November, 1964), pp. 372–84.

[22] Smaller research facilities may also lead to increased efficiency in pharmaceutical research and development. When research output is measured in terms of private rather than social gains, there appear to be substantial diseconomies of scale in research within the larger firms in this industry. See William S. Comanor, "Research and Technical Change in the Pharmaceutical Industry," *Review of Economics and Statistics* (May, 1965), pp. 182–90.

[From The Review of Economics and Statistics, Volume XLIX, November 1967, Number 4, pp. 423–440]

Advertising Market Structure and Performance

(By William S. Comanor and Thomas A. Wilson*)

This paper presents an empirical analysis of the role of advertising in consumer goods industries. The primary finding is that advertising has a statistically significant and quantitatively important impact upon profit rates which provide a measure of market performance as well as indicate the existence of market power. This result is robust, and the estimated multivariate equations account for half of the inter-industry variance of profit rates.

This finding has implications which are precisely the opposite of the conclusions reached by Telser in a recent important article.[1] This contradiction is a reflection primarily of differences in the conceptual and statistical approaches adopted rather than differences in data or sample, for with minor exceptions, we used the same set of industries, and drew upon the same basic data for advertising outlays.

We shall therefore proceed as follows. First, we shall describe the conceptual framework used. Then we shall examine the relationships which are likely to exist between product differentiation, advertising and entry barriers. Finally, we shall present the empirical results which are the core of this paper.

FRAMEWORK OF ANALYSIS

The analytical approach is to examine the joint effect of various dimensions of market structure upon profit rates. Not only do profit rates provide some indication of market performance in terms of the normal criteria of allocative efficiency, but also high returns signal the possible existence of market power.[2] If exercise in the direction of profit maximization, market power should lead to rates of return which exceed those in competitive industries that are comparable in terms of risk and growth of demand.

In this framework, concentration is simply one dimension of market structure and is not of itself a measure of monopoly or market power. Another major dimension is the height of entry barriers, which is determined in part by technical factors such as the extent of production economies of scale relative to the size of the market, the absolute amount of capital required to operate a plant of minimum efficient scale, and other absolute production cost disadvantages of new entrants.

Product differentiation, a third major dimension of market structure, plays a dual role. Not only does it directly influence the character of competition among established firms, but it also raises the height of entry barriers.[3] In this study, however, we do not deal directly with product differentiation, but focus instead upon advertising expenditures, which are both a sympton and a source of differentiation. Not only are advertising budgets influenced by product and market characteristics, but also they depend on the policies pursued by individual firms. In addition, past advertising outlays appear to be important determinant of the extent of product differentiation. Differences in advertising, therefore, reflect both structural and behavioral differences between industries.

On these grounds, the empirical analysis which follows takes the form of multi-variate regression equations which explain the inter-industry variation in profit rates as a function of different combinations of the following variables:

Seller concentration,

The rate of growth of demand,

---

*A preliminary version of this paper was presented at the December 1965 meetings of the Econometric Society. The authors thank Frank Edwards for his discussion of the paper at that time, and William Hughes and Lester Taylor for their comments on a subsequent draft. Special thanks are due Richard Caves and R. B. Heflebower, both of whom provided us with extensive and constructively critical comments.

The authors are grateful for financial assistance provided by the small grant program of the Ford Foundation and a grant from the Alfred P. Sloan Foundation.

[1] Lester Telser, "Advertising and Competition," *Journal of Poltical Economy* (Dec. 1964).

[2] Low or average profit rates do not necessarily indicate that market power is absent. Firms may become lax in minimizing costs when the discipline of competition is weak. For a discussion of such behavior, see Carl Kaysen, *U.S. vs. United Shoe Machinery Corporation* (Cambridge: Harvard University Press, 1956) 114–116.

[3] Joe Bain, *Barriers to New Competition* (Cambridge: Harvard University Press, 1956) 21,

Economies of scale in production in relation to the size of the market,
Absolute capital requirements for a plant of minimum efficient scale,[4] and
Advertising.

The specific variables used, the alternative functional forms, and other specifications of the estimated equations are described below. The conceptual relationship between advertising, product differentiation, and the height of entry barriers is discussed in the next two sections.

Before proceeding, however, it is useful to contrast the framework adopted here with that used by Telser. One of his major empirical findings is that the simple correlation between advertising outlays as a percentage of sales and the level of seller concentration is statistically insignificant. In each of the years studied, he finds that this coefficient is about 0.16, and from this, concludes that "There is little empirical support for an inverse association between advertising and competition." [5]

This approach raises the problem of whether concentration ratios are an adequate measure of the extent of competition. Telser justifies their use by stating that "Concentration of sales among the four leading firms is a widely accepted measure of monopoly." [6] While this statement is unfortunately correct, it ignores the fact that the concentration ratio measures only one dimension of market structure, and is therefore an inadequate indicator of market power, which depends on additional structural variables as well as on established behavior patterns. The significance of advertising expenditures depends on whether they represent an additional factor affecting the achievement of market power. The weak correlation between concentration and advertising simply indicates that these are independent rather than collinear variables.

ADVERTISING AND PRODUCT DIFFERENTIATION

The relationship between advertising outlays and product differentiation is important for an evaluation of the competitive effects of advertising because the former reflects the policies adopted by individual firms, while the latter is a dimension of market structure.

The degree of product differentiation in a market is measured by the cross elasticities of demand and supply which exist among competing products. Low cross elasticities of demand between these products indicate that buyers prefer the products or brands of particular sellers and will not switch in significant numbers in response to small differences in price. Low cross elasticities of supply, on the other hand, signify that firms are unable to imitate the products of their rivals sufficiently well to eliminate these consumer preferences. While cross elasticities between the products of existing producers affect the character of the rivalry which exists between them, cross elasticities between the products of established firms and potential entrants influence the height of entry barriers posed by product differentiation.[7]

Product differentiation reflects two sets of factors: the basic characteristics of products within the market, and the present and past policies of established firms with respect to advertising, product design, servicing, and distribution. On the demand side, products are more likely to be differentiable when buyers are relatively uninformed about the relative merits of existing products. This is particularly important for differentiation achieved via advertising. On the supply side, differentiation is more likely where the products of rivals cannot easily be imitated and where new entrants have difficulties in producing products which are simular to those sold by successfully established firms. In producer goods industries, successful imitation requires investment in product design and adequate service facilities. In consumer goods industries, successful imitation may require investment in advertising as well.

---

[4] No attempt was made to measure any other absolute cost disadvantages of new entrants. Bain found that only in those industries in which established firms controlled scarce natural resources were these important. Bain, *op. cit.*, 155–156. In addition, no attempt was made to measure risk. The sources of the data and various technical adjustments are described in the appendix.

[5] Telser, op. cit., 544 and 558.

[6] Ibid., 542.

[7] It is important to distinguish product differentiation from product variety. The steel industry, for example, produces a great variety of products which are sold to knowledgeable buyers, but product differentiation is minimal. In contrast, the cigarette industry offers a smaller variety of products, but product differentiation—based largely on extensive advertising—is great. Bain, op. cit., 127–129.

It is noteworthy that Bain, in his authoritative examination of product differentiation in 20 manufacturing industries, found advertising to be the most important source of product differentiation in the consumer good industries in his sample. Distribution policies are also important where forward integration in prevalent, while customer services and product design play contributing but relatively minor roles.

For typical consumer goods industries, then, a persistently high level of advertising expenditures can be viewed in two ways:

a) If firms behave reasonably, high levels of advertising indicate that the product is differentiable. In this sense, advertising is a symptom of differentiation.

b) The high level of advertising is itself an important determinant of the level of differentiation which is realized by established firms vis-à-vis potential entrants. In this sense, advertising is a source of product differentiation.

Provided that firms act reasonably, observed advertising expenditures provide a useful measure of the extent of product differentiation. We write reasonably rather than rationally since, in an oligopolistic market, rational policies are not unambiguous. What is rational policy for the group acting in concert is not rational policy for the individual firm expecting to gain a march on its rivals. It is quite possible, moreover, that rivalry via advertising among established firms is carried to the point of diminishing returns in terms of group profit rates. However, even in this case, the result of extensive advertising rivalry may be to permit the achievement of higher future profits for the group by raising entry barriers.

THE EFFECT OF ADVERTISING ON ENTRY BARRIERS

Although advertising is only one source of product differentiation, it is especially important in a number of consumer goods industries where it has a strong direct impact on entry barriers.[8] In these industries, new entrants generally are forced to sell at a price below the established brands or else incur heavy selling costs. This explains the phenomenon of unbranded products selling at prices substantially below those of highly advertised products even where there is little "real" difference between them. On this account, established firms can set prices above existing cost levels, including advertising and other selling expenses, without inducing entry.

Product differentiation via advertising affects entry barriers in three ways, each of which is analogous to the other determinants of overall entry barriers. First, high prevailing levels of advertising create additional costs for new entrants which exist at all levels of output. Because of buyer inertia and loyalty, more advertising messages per prospective customer must be supplied to induce brand switching as compared with repeat buying. Since the market which prospective entrants must penetrate is made up largely of consumers who have purchased existing products, advertising costs per customer for new entrants will be higher than those of existing firms who are maintaining existing market positions. Moreover, the costs of penetration are likely to increase as output expands and customers more inert or loyal need to be reached.[9] This effect of advertising creates an absolute cost advantage for established producers, since they need not incur penetration costs.

In addition, the effect of advertising on firm revenues is subject to economies of scale which result from the increasing effectiveness of advertising messages per unit of output as well as from decreasing costs for each advertising message purchased. The first source of economies will exist whenever the effect of advertising on consumer decisions is sufficiently important that a threshold level of advertising is required for a firm to stay in the market and maintain its current market share. In such a situation, larger firms have the advantage of being able to spread this cost over more units of output and thereby spend less per unit sold. This advantage creates economies of scale at the firm level, since an established firm does not have to spend twice as much on advertising to

---

[8] Ibid., 114–143.

[9] These penetration costs depend on past as well as current advertising outlays by established firms. The importance of past outlays is examined by Kristian S. Palda who concludes that "distributed lag models both give a better fit to the Pinkham data and forecast better than he models which do not incorporate lagged effects." *The Measurement of Cumulative Advertising Effects* (Prentice-Hall, 1964), 94.

maintain a market share which is twice that of a rival. Higher output levels are associated with lower unit costs.[10] As a result, smaller firms, including most entrants, are placed at a strong disadvantage.[11]

Economies of scale in advertising also result when the cost per advertising message declines as the number of messages supplied increases. An increased use of some forms of advertising leads to a lower most per message, and available evidence suggests that this is very important for advertising on national television and in national magazines.[12]

If advertising in a particular industry is characterized by economies of scale for either of these reasons, an entrant will suffer an additional cost disadvantage if he enters at a relatively small scale. If he enters at a scale sufficient to realize available economies of scale in advertising, however, his actions are likely to influence the price or advertising policies of the established firms. The possible reactions of established firms increase the costs and risks of entry.

Finally, if economies of scale exist either in production or in advertising, the need to obtain funds for advertising will give rise to capital requirements over and above those needed for physical plant and equipment. Furthermore, this investment in market penetration will involve a particularly risky use of funds since it does not generally create tangible assets which can be resold in the event of failure. The required rate of return on such capital will therefore be high.

These various effects are illustrated diagramatically in figure 1. Curve APC represents average production costs for established and prospective firms, and MESP is minimum efficient scale in production. Curve AAC describes average advertising costs for existing firms as well as for new entrants after they have become established. It denotes unit advertising outlays which are required in order to maintain a firm's market position and to preserve a given volume of sales once it has been established. This will depend on both the total level of advertising outlays and their distribution among established firms, and therefore, it describes prospective advertising costs for entrants only if existing firms do not react to any loss of market share. To the extent that they do respond, required advertising outlays will be higher. Curve ATC, the vertical sum of these two curves, represents average total costs for established firms.[13] MES denotes the minimum efficient scale in both production and advertising for an established firm with a given market share.

In addition, curve AMPC describes average market penetration costs for new entrants. Penetration costs represent an investment in establishing a market position and therefore depend on the opportunity cost of capital as well as on total penetration expenditures.[14]

---

[10] In the automobile industry, for example, the two smaller firms during the 1950's were forced to spend more than twice as much on advertising per car sold as did either Ford or General Motors. Between 1954 and 1957 Studebaker and American Motors spent annually on national advertising approximately $64.04 and $57.89, respectively, per automobile sold while G.M. spent $26.56 per unit and Ford spent $27.22 per unit. Chrysler was in an intermediate position, spending $47.76 per unit. Leonard W. Weiss, *Economics and American Industry* (New York: John Wiley and Sons, 1961) 342.

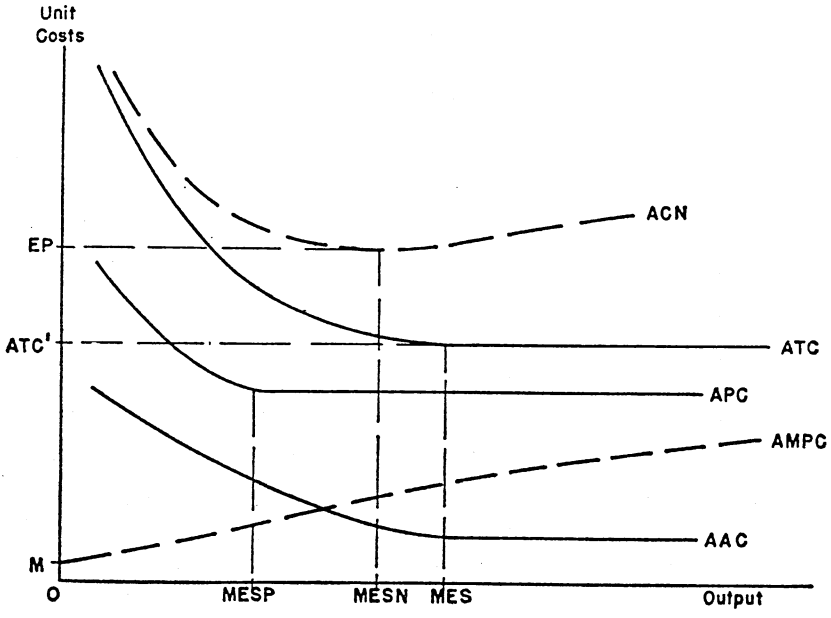
[11] This result occurs *within* the relevant market. When a firm in a regionally segmented market expands its national market share by moving into new geographic areas, unit advertising costs do not decline.

[12] The extent of discounts given to large advertisers is documented in *Federal Trade Commission vs. The Procter & Gamble Company,* Brief for the Federal Trade Commission in the Supreme Court of the United States (Dec. 1966), 12–13.

[13] For simplicity, we assume here that advertising constitutes the only form of selling expense.

[14] Penetration costs include extra advertising outlays which are required for entry. These outlays will represent total penetration costs if the price charged by the entrant is the same as that set by established producers. If the entrant is forced to set a price below that of existing firms, there are additional penetration costs which equal the price differential times the amount of output sold by the entrant at the lower price.

FIGURE 1. — ADVERTISING AND ENTRY BARRIERS



This schedule therefore denotes the required rate of return on capital invested in market penetration times the total expenditure required to establish a given volume of sales, all divided by the number of units sold. The figure illustrates the case where average penetration costs rise throughout the relevant range of output. This assumes that the growing difficulty involved in winning over customers with stronger preferences for the products of established firms, reinforced by rising required rates of return as the absolute amount of capital required for penetration increases with the scale of entry, is not fully offset by economies of scale in advertising or by bandwagon effects for the new entrant's products.

Curve ACN represents average costs, including peneration costs, for new entrants, and MESN is the most efficient scale for entry if the reactions of established producers are neglected. From this, it follows that EP is the *minimum* price at which entry will occur. If MESN is a negligible fraction of the market, EP is the entry-inducing price. If, however, MESN is a significant fraction of the market, entry is unlikely to occur even at price EP because the entrant will expect established producers to contest the encroachment of their market position through an increase in advertising outlays or by a reduction in price. The gap between EP and ATC′ represents, therefore, the minimum price-cost margin which may induce entry.

This figure demonstrates, moreover, that the interaction between rising penetration costs and economies of scale at the firm level is important even if no allowance is made for the reactions of existing producers. If economies of scale in both production and in advertising were absent, the relevant price-cost margin would be simply M, which is less than EP–ATC′.

VARIABLES USED IN THE MULTIVARIATE ANALYSIS

In this section, we brefly define and discuss the rationale for the selection of each of the specific measures.

Profit Rates

The profit rate variable used is profits after taxes as a percentage of stockholders' equity,[15] averaged within each industry for firms with assets exceeding

[15] This profit rate variable was used originally in Joe S. Bain, "Relation of Profit Rates to Industry Concentration: American Manufacturing, 1936–1940," *Quarterly Journal of Economics* (Aug. 1951), 296–297, and Bain, *Barriers to New Competition*, 192.

$500,000. This procedure avoids the difficulty, noted by Stigler, of profit withdrawals in the form of executive salaries in small and closely held corporations.[16] Profit rates are also averaged for the period 1954–1967, which covers a complete business cycle.

Although the profit rate on stockholders' equity is viewed as a more appropriate variable than the rate of return (including interest) on total assets,[17] we examined whether the empirical results would be sensitive to this decision. The simple correlation between the profit rate on stockholders' equity and the rate of return on total assets is 0.93. In addition, the correlation coefficients between each of these variables and the 1954 four-firm concentration ratio[18] are, respectively, 0.36 and 0.33. These results suggest that our empirical findings are unlikely to be sensitive to the choice of a specific profit rate valuable.

Advertising

In light of the discussion in the preceding section, it is useful to examine the absolute volume of advertising expenditures by existing firms as well as the advertising-sales ratio. The latter variable probably provides a good indication of the absolute cost disadvantage of the new entrant at small scales of entry, but is likely to be a less accurate index of the economies of scale and absolute capital requirements effects of advertising.

We have, therefore, calculated two measures of advertising intensity: advertising outlays per dollar of sales for firms with assets greater than $500,000, and average advertising expenditures per firm among firms which account for 50 per cent of industry output.[19] Both advertising variables are averages for the years 1954 through 1957.

Seller Concentration

Concentration is a sufficiently prominent variable in the literature to warrant introducing it in three alternative formulations. First, a trichotomous classification based on Kaysen and Turner's classification of market groups is used.[20] Second, the average four-firm concentration ratios published by Stigler are introduced.[21] Finally, a dichotomous classification is constructed on the basis of Bain's finding that a critical point is reached when the eight-firm concentration ratio exceeds 70 per cent.[22]

Economies of Scale in Production

Economies of scale in production presumably exist primarily at the level of the plant rather than the firm. In the absence of better estimates for most of the industries in the sample, a measure is derived from the size distribution of plants within the relevant industries. Since cost minimization is an element of profit maximization, large multi-plant firms should operate plants which are sufficiently large to realize available scale economies. Where demand is not a limiting factor, moreover, competition among firms should lead directly to plants which equal or exceed minimum efficient scale.[23] At the same time, however, small plants may exist. These may have been built in an earlier period, before demand had expanded or a technology which required large scale had been developed, or

---

[16] George J. Stigler, *Capital and Rates of Return in Manufacturing Industries* (Princeton University Press) 125–127.

[17] This is because firms presumably maximize profits, rather than the sum of profits plus interest payments. The rate of return on stockholders' equity will therefore be a more sensitive indicator of the extent of freedom from competitive constraints.

[18] In two cases out of 41, our industry classifications differed from those presented by Stigler. In both of these cases, "Screens and Venetian Blinds," and "Radio, T.V., and Phonograph," it appeared that Stigler had combined these industries with smaller, miscellaneous industries. In these calculations, therefore, we used data for the more aggregated industry to stand for its major component.

[19] The procedure used was to select successive asset size classes of firms until 50 percent or more of industry sales was covered. The proportion of sales in the boundary size class required to reach this degree of coverage was used to determine the amount of advertising and the number of firms from that size class included in both the numerator and denominator of the measure of advertising per firm. For some industries, the largest size class accounted for more than 50 per cent of sales. In such cases, the measure is simply advertising per firm in the largest size class.

[20] Carl Kaysen and Donald F. Turner, *Antitrust Policy*, 27.

[21] Stigler, *op. cit.*, 214–215.

[22] Bain, "Relation of Profit Rates to Industry Concentration," 314.

[23] Since the bulk of the evidence suggests that cost curves in manufacturing are "L shaped" rather than "U shaped," plants which exceed minimal efficient scale will typically be efficient plants. See J. Johnston, *Statistical Cost Analysis* (New York: McGraw-Hill, 1960) 44–168.

they may result from the entry of small firms. They may also exist in pockets of the market which are geographically segmented or may specialize in narrow product lines which are not representative of the industry generally. It is therefore important to select a measure which is insensitive to the entry of single-plant firms of sub-optimal scale.

The measure used is based on average plant size among the largest plants accounting for 50 per cent of industry output. This average plant size is divided by total output in the relevant market to obtain the scale economies variable used in the regression analyses.[24]

A test of the reliability of this variable can be made by comparing minimum efficient scale as a percentage of industry output with Bain's estimates. Not only did Bain concentrate on a smaller number of industries, but also he used varied forms of information. Therefore, his estimates can be considered a benchmark against which to appraise various methods of estimating the extent of scale economies. We examine both the method described above and an alternative method, the Survival Technique, as used by Weiss.[25]

Of the 20 industries examined by Bain, data on the size distribution of plants are available for 19. Across those industries, the correlation coefficient between Bain's estimates and those derived from the method proposed above is 0.89.[26] Estimates based on the Survival Technique are available for 13 of the industries studied by Bain, and the correlation between these estimates and those presented by Bain is 0.66. When the comparison is limited to these same 13 industries, the correlation coefficient between estimates derived from the size distribution of establishments and those published by Bain is 0.86. The method proposed above is more consistent with Bain's estimates than those computed from the Survival Technique,[27] and it will therefore be used in the succeeding analysis.[28]

Absolute Capital Requirements

This amount of capital required for entry at the scale of a single efficient plant is based upon the above estimates of economies of scale. The average output level of plants at estimated minimum efficient scale is multiplied by the ratio of total assets to gross sales for the industry.[29]

The Rate of Growth of Demand

The rate of growth of demand is measured by the rate of growth of sales between 1947 and 1957.[30] A period of this length was chosen in order to emphasize the long-run effects of the growth of demand, and the terminal years selected were both years of nearly full employment for the economy as a whole.

Composite Variables Representing Technical Entry Barriers and Advertising

Industries were classified into three groups on the basis of the two variables which measure technical entryl barriers (economies of scale and absolute capital requirements). Dummy variables identifying industries with high and moderate technical entry barriers were used in some regression models.[31]

---

[24] An alternative measure based on average plant size among the largest plants accounting for 70 per cent of output was also constructed. This variable was highly correlated with the variable used.

[25] See Leonard Weiss, "The Survival Technique and the Extent of Suboptimal Capacity," *Journal of Political Economy*, (June 1964), 246–261.

[26] In these calculations, Bain's estimates of minimum efficient scale for the Steel Industry refer to "Steel Works and Rolling Mills," while for the Copper Industry, they refer to "Primary Copper."

[27] In addition, the survivor technique has the great disadvantage of frequently yielding indeterminate results. See Weiss, op. cit., 258–259.

[28] Ideally, some account should be taken of the slope of the average cost function below minimal efficient scale. The variable used in this paper, in contrast to the estimates developed by Bain, ignores this aspect of economies of scale.

[29] This measure is likely on the average to understate capital requirements. The book value of total assets will normally be less than their replacement cost, as a result of inflation in preceding years. In addition, a new firm is likely to have higher input costs while it is learning the production and distribution techniques required in the market.

[30] The ideal measure, of course, would be the rate at which the demand curve shifted over time. The rate of growth of sales is an exact measure only if the price elasticity of demand is unity or if prices did not change over the period.

[31] High technical entry barriers are assumed to exist if either the scale of an efficient plant exceeded ten per cent of the market or the capital required for efficient entry equalled at least $50 million. In addition, if scale economies fell between six and 9.9 per cent *and* if capital requirements amounted to more than $25 million, the same classification is designated. Low technical barriers exist when either scale economies amounted to less than three per cent and capital requirements are less than $25 million *or* when the former was less than six per cent but the latter sum was under $10 million. Moderate technical barriers are assumed in all other cases.

Similarly, industries were classified into three groups on the basis of the two advertising variables. Dummy variables identifying industries with high and moderate advertising expenditures were also used as an alternative measure of advertising intensity in some equations.[32]

Local Market Dummy Variable

In some of the regression equations, a dummy variable was introduced to identify the three local market industries.

THE SAMPLE OF INDUSTRIES

For the reasons presented above, the empirical analysis is confined to consumer goods industries.[33] Of the 41 industries included in the analysis, 29 produced non-durable consumer goods, and the remaining 12 produced consumer durables. In size, the industries ranged from motor vehicles with average sales of over $20 billion per year to hats with average sales of only $122 million per year. Spurious size effects are absent, however, since the dependent variable is expressed in ratio form. The core of the analysis is based upon unweighted interindustry regressions. However, as subsequent tests indicate that heteroscedasticity is present, weighted regressions are also estimated.

Three of the industries—soft drinks, bakery products, and dairy products—sell in local markets, and this factor influences the appropriate measures of both concentration and scale economies. Two techniques are used to handle this problem:

a) Economies of scale are estimated in relation to the typical local market, and the Kaysen and Turner concentration classifications which take into account the local market character of these industries, are used.

b) In some equations, national concentration ratios are introduced along with a dummy variable which identifies the local market character of these three industries.

One interesting characteristic of the underlying data is that the distribution of advertising-sales ratios across the 41 industries is highly skewed. Twenty-five of the industries have ratios below three per cent while eight have ratios between three and six per cent and only six have ratios which exceed six per cent. In the latter group, perfumes have an advertising-sales ratio of 15 per cent, cereals and drugs ten per cent each, soap nine per cent, malt liquor seven per cent, and soft drinks slightly over six per cent. Notable industries in the intermediate group include cigarettes and wines with ratios of about five per cent each.

Although advertising per firm is positively correlated with advertising per dollar of sales, the positions of two important industries change radically depending on which variable is used to measure advertising intensity. While tires and tubes and motor vehicles have quite low advertising-sales ratios, both are among the small group of industries with high or very high average advertising outlays per firm among the leading firms. As a result, tires and tubes and motor vehicles are classified respectively as industries with moderate and high overall advertising barriers.

MAJOR EMPIRICAL RESULTS

The simple correlation coefficients between profit rates and each explanatory variable are presented in table 1. All of the coefficients have the expected sign and all are statistically significant in at least one functional form. Moreover, the lograrithmic relationship appears appropriate in the case of the growth of demand variable, while the opposite is the case with regard to the advertising-sales ratio. As the latter variable is already expressed as a percentage, it is measured in units comparable to the dependent variable.

---

[32] High barriers are assumed either when the advertising-sales ratio exceeded eight per cent or when advertising expenditures among leading firms averaged more than $20 million. The same classification is given to industries where the average ratio fell between four and eight per cent *and* average annual expenditures amounted to between $5 and $20 million. Low barriers exist when the advertising-sales ratio was less than two per cent and average expenditures less than $5 million *or* when the ratio lay between two and four per cent and expenditures did not exceed $1 million per firm. In other cases, moderate advertising barriers are assumed to exist.

[33] The Petroleum industry is excluded because of the difficulty of obtaining profit data comparable to other industries in view of the special tax treatment of that industry.

TABLE 1.—*Simple correlation coefficients—Profit rates and various dimensions of market structure*

Market structure variables	Correlation with profit rates	
	Natural units [1]	Logarithm [2]
Growth of demand	0.17	0.42
Capital requirements	[3] 0.43	[3] 0.57
Economies of scale	0.25	[3] 0.37
Advertising-sales ratio	[3] 0.42	[4] 0.27
Advertising per firm	[3] 0.43	[3] 0.50
Concentration ratio (4 firms)	[4] 0.36	[4] 0.35

[1] The units of measurement are described in the text above.
[2] In computing these coefficients, the structural dimension is measured in logarithms although the profit rate is measured in natural units.
[3] Indicates coefficient is statistically significant at the 99 percent level.
[4] Indicates coefficient is statistically significant at the 95 percent level.

NOTE.—Tests of significance are made on the basis of 1-tailed t tests.

The correlation between profit rates and concentration in table 1 is based on a continuous four-firm concentration ratio. The relationship between these two variables was also examined in terms of discrete groupings. Industries were divided according to both the three-way classification scheme proposed by Kaysen and Turner, and a two-way classification depending on whether the eight-firm concentration ratio exceeded or was less than 70 percent. The following results are obtained:

	Number of industries	Average profit rates
Kaysen and Turner trichotomy:		
Type I oligopolies	13	8.4
Type II oligopolies	14	9.2
Unconcentrated	14	6.3
Dichotomy based on 8-firm concentration ratio at 70 percent:		
Concentrated	8	10.0
Unconcentrated	33	7.5
All industries	41	7.9

While the distinction between Type I and Type II Oligopolies seems, to the average, unimportant, there do appear to be substantial differences in profits between concentrated and unconcentrated industries. These differences are important in both classification schemes.

The core of the empirical work is the multiple regression equations which relate profit rates to various combinations of the explanatory variables. A set of linear equations is presented in table 2. As may be observed, the advertising-sales ratio and the measure of capital requirements appear to be the most important explanatory factors. Their regression coefficients are generally significant even when all other variables are included. The variable describing economies of scale seems from these results to be quite weak, although it has the expected sign in all cases. The advertising per firm coefficient is significant if the advertising-sales ratio is not included. Where both are included, it tends to be insignificant. Advertising outlays per firms are correlated with absolute capital requirements (the simple correlation coefficient between these variables is 0.40) and this variable is not as statistically important in conjunction with the latter as is the advertising-sales ratio.

TABLE 2.—*Multiple regression equations explaining profit rates—Linear results*

	Intercept	Advertising-sales ratio	Advertising per firm	Economies of scale	Absolute capital requirements	Growth of demand	Concentration classses[a] Type I	Type II	R^2	Corrected R^2
(1)	0.049	*0.424 (2.4)	0.00000059 (1.1)	0.113 (0.8)	**0.000281 (3.0)	0.0014 (0.6)	−0.0158 (1.0)	0.0084 (0.6)	**0.47	**0.34
(2)	0.052	0.296 (1.6)	*0.00000114 (2.0)	0.113 (0.7)		0.0018 (0.7)	−0.0058 (0.3)	0.0115 (0.8)	*−0.32	*0.19
(3)	0.051	**0.437 (2.5)	0.00000060 (1.1)		**0.000282 (3.0)	0.0012 (0.5)	−0.0091 (0.7)	0.0126 (1.0)	**0.46	**0.35
(4)	0.048	**0.499 (3.1)		0.116 (0.8)	**0.000315 (3.5)	0.0016 (0.7)	−0.0117 (0.7)	0.0085 (0.6)	**0.45	**0.34
(5)	0.058		*0.00000112 (2.1)	0.145 (0.9)	*0.000227 (2.3)	0.0012 (0.5)	−0.0147 (0.9)	0.0146 (1.0)	**0.38	**0.25

[a] Based on Kaysen and Turner groupings.

NOTE.—Figures in parentheses are t values. The statistical significance of the regression coefficients is tested by means of 1-tailed t test and of the multiple correlation coefficients by means of the F-ratio test.
*Indicates coefficient is statistically significant at the 95-percent level.
**Indicates coefficient is statistically significant at the 99-percent level.

In none of the equations do the estimated coefficients of the concentration dummy variables exceed their standard errors. In addition, the coefficient for Type I Oligopolies has a negative sign throughout, which does not coincide with a priori expectations. While the impact of concentration is examined at greater length below, the linear results suggest that the partial effect of this variable may be relatively unimportant when it is introduced in conjunction with variables reflecting product differentiation, the height of technical entry barriers, and the rate of growth of demand.

In an alternative formulation, four dummy varaibles were defined to represent high and moderate technical entry barriers and high and moderate advertising intensities, and these were introduced in place of the advertising, economies of scale, and absolute capital requirements variables. The results are presented in table 3. The dummy variables designed to measure the influence of technical barriers are not statistically significant, and in the second equation, the estimated coefficient for industries with high technical barriers is less than the coefficient for industries with moderate barriers. This result reflects in part the correlation between the high technical barrier dummy variable and concentration. The correlation coefficient between the dummy variable for high technical barriers and the Type I concentration dummy variable is 0.53. When the four-firm concentration ratio is used, this coefficient rises to 0.68. This collinearity obscures the separate effects of concentration and technical entry barriers.

TABLE 3.—*Multiple regression equations containing composite variables*

	Intercept	Advertising barriers		Technical entry barriers		Concentration [a]		Growth of demand (logs)	R^2	Corrected R^2
		High	Moderate	High	Moderate	Type I	Type II			
(1)	0.056	**0.0552 (3.3)	0.0080 (0.6)	0.0154 (0.9)	0.0090 (0.5)	−0.0132 (0.7)	−0.00085 (0.05)	*0.020 (1.9)	**0.46	**0.33

	Intercept	Advertising barriers		Technical entry barriers		Concentration ratio [b]	Regional industry dummy variable	Growth of demand (logs)	R^2	Corrected R^2
		High	Moderate	High	Moderate					
(2)	0.044	*0.0403 (2.3)	−0.0049 (0.3)	−0.0150 (0.8)	−0.0013 (0.1)	0.000596 (1.4)	0.0311 (1.4)	*0.024 (2.4)	**0.48	**0.36

[a] Based on Kaysen and Turner groupings.
[b] Four-firm concentration ratio.

NOTE.—Figures in parentheses are t values. The statistical significance of the regression coefficients is tested by means of one-tailed t test and of the multiple correlation coefficients by means of the F-ratio test.
*Indicates coefficient is statistically significant at the 95-percent level.
**Indicates coefficient is statistically significant at the 99-percent level.

To consider further the impact of concentration, the continuous concentration ratio was introduced into the analysis. As this variable is available only on a national basis, a dummy variable was also introduced to identify industries which sell in local rather than in national markets.[34] It is interesting that the continuous variable appears to have a stronger impact on profits than do the concentration dummy variables. Both the coefficient for concentration and for the local market dummy variable are significant at the 90 per level.

In contrast to the weak effect of composite technical barriers, the composite advertising variable has a strong effect. In both equations, the high advertising variable is statistically significant and the coefficient appears relatively stable. In addition, the growth of demand variable becomes significant when introduced in logarithmic form.

Additional sets of regression equations are presented in tables 4 and 5. In both sets, capital requirements are introduced in logarithmic form and the estimated coefficients are generally significant. However, some degree of collinearity exists between the measures of capital requirements and scale economies. When both variables are introduced into the equations, the variance of the estimates increases, and in some cases, the capital requirements coefficient becomes statistically insignificant.

In table 4, the advertising-sales ratio is used to measure the entry barriers created by high advertising expenditures, while in table 5, the dummy variable representing high advertising barriers is introduced. Both advertising variables are statistically significant in all of the equations presented, and the estimated coefficients of both variables are very stable. The dummy variable for high advertising barriers appears to be somewhat stronger than the advertising-sales ratio. However, the coefficients of the other structural variables included in these equations are not sensitive to the particular advertising variable used.

TABLE 4.—*Multiple regression equations explaining profits rates—Major findings with advertising-sales ratio*

	Intercept	Advertising-sales ratio	Capital requirements (logs)	Economies of scale (logs)	Growth of demand (logs)	Concentration class [a]	Regional industry dummy variable	R^2	Corrected R^2
(1)	0.042	*0.362 (2.4)	**0.0097 (3.2)		0.016 (1.6)			**0.46	**0.40
(2)	0.042	*0.362 (2.3)	**0.0096 (2.5)	0.000067 (0.01)	0.016 (1.6)			**0.46	**0.38
(3)	0.039	*0.343 (2.3)	**0.0105 (2.8)		0.015 (1.4)	0.0043 (0.3)	0.0278 (1.5)	**0.49	**0.40
(4)	0.038	*0.341 (2.3)	**0.0111 (3.6)		0.014 (1.4)		0.0280 (1.6)	**0.49	**0.42

[a] An industry is concentrated if the 8-firm concentration ratio equals or exceeds 70 percent; otherwise it is unconcentrated.

*Indicates coefficient is statistically significant at the 95-percent level.

**Indicates coefficient is statistically significant at the 99-percent level.

NOTE.—Figures in parentheses are t values. The statistical significance of the regression coefficients is tested by means of one-tailed t test and of the multiple correlation coefficients by means of the F-ratio test

[34] These market characteristics had already been accounted for in the Kaysen and Turner groupings.

TABLE 5.—*Multiple regression equations explaining profit rates—Major findings with high advertising barrier*

	Intercept	High advertising barrier	Capital requirements (logs)	Economies of scale (logs)	Growth of demand (logs)	Concentration class [a]	Regional industry dummy variable	R^2	Corrected R^2
(1)	0.053	**0.0379	*0.0066		*0.018			**0.48	**0.42
		(2.8)	(2.0)		(1.9)				
(2)	0.069	**0.0388	0.0047	0.0038	*0.019			**0.49	**0.42
		(2.8)	(1.1)	(0.7)	(2.0)				
(3)	0.048	**0.0395	*0.0089		0.015	−0.0063	*0.0318	**0.53	**0.45
		(2.9)	(2.4)		(1.5)	(0.5)	(1.8)		
(4)	0.048	**0.0379	**0.0082		*0.016		*0.0316	**0.52	**0.46
		(2.9)	(2.5)		(1.7)		(1.8)		

[a] An industry is concentrated if the 8-firm concentration ratio equals or exceeds 70 percent; otherwise it is unconcentrated.

NOTE.—Figures in parentheses are t values. The statistical significance of the regression coefficients is tested by means of 1-tailed t test and of the multiple correlation coefficients by means of the F-ratio test.
*Indicates coefficient is statistically significant at the 95 percent level.
**Indicates coefficient is statistically significant at the 99 percent level.

In both tables 4 and 5, a dummy variable identifying industries with eight-firm concentration ratios which exceed 70 per cent was introduced in order to examine the effect of this aspect of market structure in yet another specification. The coefficients, however, remain smaller than their standard errors.

The regional industry dummy variable was also introduced into the equations in both tables. In all cases, its estimated parameters are significant at the 90 per cent level, and reach the 95 per cent level in table 5. While this variable was used originally to correct for the use of concentration ratios calculated on a national basis, it appears to have an independent effect which does not depend on the presence of the other variable. It is useful, therefore, to compare the structural features of the three local market industries included in our sample with the others. Relevant data are presented in table 6. The scale economy variable is the most sensitive to this industry characteristic. As would be expected, these industries, on the average, have much higher estimates of the ratio of minimum efficient scale to market than do the national industries. We should expect, therefore, that the local market dummy variable represents the increased importance of economies of scale as well as the higher concentration levels in local markets.

The multiple correlation coefficients for these equations are aways statistically significant. The included variables typically account for about half of the total variation in industry profit rates.

TABLE 6.—*Local and regional industry characteristics*

Industry	Concentration class [1]	Advertising-sales ratio (percent)	Advertising per firm (millions)	Capital requirements (millions)	Economies of scale (percent)	Growth of demand (ratio)	Profit rate (percent)
Soft drinks	Type I	6.2	$0.26	$0.75	8.2	1.98	10.0
Dairy	Type II	2.2	15.12	2.09	14.2	1.16	7.9
Bakery	Type II	2.9	1.97	2.57	8.3	1.69	9.3
Average of 3 industries		3.8	5.78	1.80	10.2	1.61	9.1
All industries		3.3	6.03	24.32	4.7	1.83	7.9

[1] Kaysen and Turner groupings.

THE PROBLEM OF MULTICOLLINEARITY

As was noted above, a number of the explanatory variables included in the analysis are collinear to some extent. While the simple correlations between the advertising-sales ratio and the other independent variables are typically low,[35] it is useful to examine the sensitivity of the estimated coefficients for this variable to changes in the specifications of the regression equations. The results are presented in table 7. The estimated coefficients are reasonably stable, ranging from 0.30 to 0.52. In addition, seven of the 15 coefficients presented are significant at the 99 per cent level, seven at the 95 per cent level, and the remaining coefficient at the 90 per cent level.

TABLE 7.—*Sensitivity of regression coefficients for advertising variables to changes in specification of regression equations*

A. ADVERTISING-SALES RATIO

[N—variable in natural units; L—variable in logarithms; I and II denote Kaysen and Turner concentration classes; H and M denote high and moderate dummy variables]

Estimated coefficient	Value of t	Concentration	Regional industry dummy	Capital requirements	Scale economies	Growth of demand	Advertising per firm
0.42	2.27	I, II	----	----	N	----	----
0.30	1.61	I, II	----	----	----	----	N
0.42	2.40	I, II	----	N	N	----	N
0.43	2.50	I, II	----	N	----	----	----
0.48	2.75	----	----	N	N	----	N
0.52	3.24	----	----	N	N	N	----
0.50	2.92	N	N	----	N	----	----
0.34	2.21	I	----	L	N	L	----
0.35	2.29	I	----	L	N	L	----
0.49	3.28	----	----	N	N	L	----
0.41	2.41	----	----	N	N	L	N
0.40	2.54	N	N	----	----	L	----
0.46	3.02	N	N	----	----	L	----
0.36	2.43	----	----	----	----	L	----
0.33	2.19	N	N	----	----	L	----

B. DUMMY VARIABLES FOR HIGH ADVERTISING BARRIERS

Estimated coefficient	Value of t	Other variables included						
		Concentration	Regional industry dummy	Capital requirements	Economies of scale	Growth of demand	Technical barriers	Moderate advertising barrier
0.040	2.28	N	N	----	----	L	H, M	M
0.049	3.67	I, II	----	----	----	L	----	----
0.050	3.25	I, II	N	----	----	L	H, M	M
0.045	3.39	N	N	----	----	L	----	----
0.038	2.80	----	----	L	----	L	----	----
0.039	2.83	----	----	L	L	L	----	----
0.038	2.89	N	N	L	----	L	----	----

The stability of the estimated coefficient for the dummy variable denoting high advertising barriers is also tabulated. These coefficients appear to be insensitive to changes in the specifications of the equations. They lie between 0.038 and 0.050. Six of the seven coefficients presented are significant at the 99 per cent level and the remaining one at the 95 per cent level. The estimated effect of either the advertising-sales ratio or the high advertising dummy therefore does not appear to be affected by which of the other variables are included in the equations.

[35] The simple correlation coefficients between the advertising-sales ratio and other structural variables are as follows: the log of economies of scale, 0.27; the log of capital requirements, 0.21; the log of growth of demand, 0.40; and the four-firm concentration ratio, 0.10.

While there may be other variables which affect the estimated relationships between advertising and profits, the importance of both advertising variables is relatively insensitive to changes in specification of the variables and the models examined in this paper.[36]

Significant correlations also exist between capital requirements, economies of scale, and concentration. This is not surprising, for one should expect the first two variabes to have some effect on the latter. At the same time, however, concentration is influenced by other factors, such as the past record of merger activity in the industry. To examine the extent to which concentration is explained by scale economies and capital requirements, two multiple regression equations were fitted. The results, which are striking, appear in table 8.

TABLE 8.—*Multiple regression analysis—Concentration and technical entry barriers*

Concentration [a]	Intercept	Capital requirements (logs)	Economies of scale (logs)	Regional industry dummy variable	R^2	Corrected R^2
(1) Natural units.....	49.9	**7.08 (3.9)	**6.91 (2.6)	−11.2 (1.2)	**0.71	**0.68
(2) Logarithms.....	3.85	**0.244 (5.1)	**0.238 (3.4)	−0.294 (1.2)	**0.81	**0.79

[a] Four-firm concentration ratios.
Figures in parentheses are t values. The statistical significance of the regression coefficients is tested by means of one-tailed t test and of the multiple correlation coefficients by means of the F-ratio test.
**Indicates coefficient is statistically significant at the 99-percent level.

Absolute capital requirements, scale economies, and the local market dummy variable [37] together account for a substantial share in the variation in national concentration ratios. In logarithmic form, over 80 per cent of the variation is explained by these variables. What is surprising is the small share of variation left to be accounted for by other factors. With this high a degree of inter-correlation, it is understandable that the estimated coefficients for concentration are not statistically significant. The role of cencenration appears closely linked to that of technical entry barriers and there is little remaining influence which is evident.[38]

HETEROSCEDASTICITY AND WEIGHTED REGRESSIONS

An examination of the residuals from a leading equation (number 4 in table 5) revealed that heteroscedasticity is present, as small industries typically have large residuals. There are two possible reasons for this phenomenon. The smaller industries may tend to have fewer firms, so that the variance of average profit rates is larger. The smaller industries may also have smaller firms. Previous studies have indicated that the variance of profit rates among small firms is greater than among larger firms,[39] and this would also account for a larger variance for smaller industries.

[36] This result, however, does not apply to average advertising expenditures per firm, which is more strongly correlated with the other explanatory factors. As a result, its statistical significance in regression analysis appears to depend on which of the other variables are included in the estimating equation.

[37] The regional industry dummy variable was included because the concentration ratios are constructed on a national basis. The negative sign on the coefficient represents simply the downward bias of the national ratios in those industries.

[38] One should be wary of drawing any policy conclusion on the basis of this equation. Merger activity may be highly correlated with entry barriers. Furthermore, there is some element of spurious correlation between the scale economies measure and concentration. The scale economies measure used here is 0.5 times the reciprocal of the number of the largest plants required to account for one-half of industry output. It is, therefore, related to plant concentration. Since plant concentration and firm concentration may be expected to be correlated even in the absence of variations in relative scale economies, some spurious correlation exists between concentration and relative scale economies. (The authors are indebted to Joe S. Bain for the elaboration of this point.)

[39] Sydney S. Alexander, "The Effect of Size of Manufacturing Corporation on the Distribution of the Rate of Return, this REVIEW, XXXI (Aug. 1949), 229–235.

To determine an appropriate weighting scheme, an empirical approach was adopted. The variance of the residuals was calculated for successive quartiles in the distribution of industry sales. From this tabulation, it was clear that the use of industry sales is inappropriate as a weighting variable as it would give too much emphasis to the largest industries. The square root of sales, however, is nearly proportional to the variance of the residuals, and was therefore chosen as the weighting variable.

The weighted regressions were fitted both for all industries, and for all industries except motor vehicles, as this industry is an outlying observation with respect to some of the variables, including the weighting variable. The results appear in tables 9 and 10. As is clear, the R^2 of each of the weighted regressions is considerably higher than the R^2 of its unweighted counterpart. This is to be expected, since the weighting procedure deliberately emphasizes industries with smaller residuals and the R^2 measures the proportion of the weighted variance of the dependent variable explained by the regression equation. Another way of looking at this is that weighting essentially involves multiplying the equation by the root of the weights (in this case by the fourth root of sales) and proceeding by ordinary least squares. The R^2 indicates the success at predicting profit rates multiplied by the fourth root of sales.[40]

The results are impressive. About 75 percent of the weighted variance across all industries is accounted for by these equations and about 65 percent of the weighted variance is explained when the outlying auto industry is excluded.

TABLE 9.—*Weighted regressions with advertising sales ratio*

	Intercept	High advertising-sales ratio	Capital requirements (logs)	Economies of scale (logs)	Growth of demand (logs)	Concentration ratio	Regional industry dummy variable	R^2	Corrected R^2
(1) a. All industries	0.040	*0.29 (1.9)	**0.013 (4.8)		0.0084 (1.0)		**0.028 (1.9)	**0.76	**0.72
b. Motor vehicles excluded	0.045	**0.44 (3.1)	*0.0077 (2.4)		0.0096 (1.3)		0.020 (1.4)	**0.67	**0.62
(2) a. All industries	0.066	*0.28 (1.8)	**0.010 (3.1)	0.0046 (0.9)	0.0096 (1.1)			**0.75	**0.71
b. Motor vehicles excluded	0.074	**0.42 (2.8)	0.0040 (1.2)	0.0052 (1.2)	0.011 (1.4)			**0.67	**0.62
(3) a. All industries	0.040	*0.29 (1.9)	*0.014 (2.4)		0.0081 (0.9)	−0.00003 (0.08)	*0.028 (1.8)	**0.76	**0.72
b. Motor vehicles excluded	0.047	**0.43 (2.9)	0.0090 (1.6)		0.0088 (1.1)	−0.00011 (0.3)	0.021 (1.4)	**0.67	**0.61

* Indicates coefficient is statistically significant at the 95-percent level.
** Indicates coefficient is statistically significant at the 99-percent level.

NOTE.—Figures in parentheses are t values.

The high advertising barrier dummy variable and the advertising-sales ratio variable are introduced alternatively. The former is significant at the 99 per cent level in all equations, the latter at the 95 per cent level when the auto industry is included, at the 99 per cent level otherwise. The collinearity between capital requirements, economices of scale and concentration is again evident, but in contrast to the unweighted regressions, the economies of scale variable is sometimes significant when introduced alongside capital requirements (and the latter variable is sometimes insignificant).

[40] It is important to note that the increase in R^2 is no indication that the weighting used is the correct one. Indeed, a very high R^2 can be obtained by weighting with industry sales, which is clearly inappropriate. A subsequent test, moreover, was made on the extent of heteroscedasticity in the weighted regressions. The residuals from equation 1a in table 10 were calculated and the successive variances of these residuals were compared with the mean root sales in the relevant quartile. The fact that the two variables were nearly proportional provides some confirmation of the use of the square root of sales as the weighting variable.

These results, moreover, provide additional evidence of the stability of the coefficients for these two advertising variables. They also point to the joint significance of technical barriers to entry and of concentration, but the collinearity among these variables prevents precise measurement of the separate effects of each of these variables.

THE PROBLEM OF CAUSALITY

We have found that the inter-industry variation in profit rates can be explained quite well by a model incorporating the rate of growth of demand, some measure of advertising intensity, and variables reflecting the importance of concentration and technical barriers to entry. The relationship between profits and either of the advertising variables introduced into the equations is quite robust. Throughout this paper, we have assumed that the direction of causality is from the independent variables to profit rates. Could the reverse be the case?

A plausible case can be made that a significant feedback exists from profits to advertising expenditures, since advertising reflects the discretionary behavior of firms as well as the extent of product differentiation. Indeed, we should not be surprised if a time-series analysis, which emphasizes short-run effects, revealed that changes in profits preceded, rather than followed, changes in advertising expenditures.

There are a number of factors, however, which suggest that the causality of the observed relationships runs largely from advertising expenditures to profits. A cross-sectional study tends to emphasize the long-run differences between industries, and this in turn is more likely to reflect the structural rather than the behavioral aspects of advertising. Profit levels cannot influence those market and product characteristics which permit product differentiation via advertising. Firms with high profit rates will not have higher *optimum* advertising expenditures than firms with low profit rates in the same market situation. The pursuit of profits will hence limit the extent to which profits will be "spent" on advertising, especially over a period of several years.

TABLE 10.—*Weighted regressions with high advertising barrier*

	Intercept	High advertising barrier	Capital requirements (logs)	Economies of scale (logs)	Growth of demand (logs)	Concentration ratio	Regional industry dummy variable	R^2	Corrected R^2
(1) a. All industries	0.052	**0.035 (2.9)	**0.0080 (2.5)	----------	0.012 (1.5)	----------	*0.027 (1.9)	**0.78	**0.75
b. Motor vehicles excluded	0.055	**0.032 (2.7)	*0.0064 (1.8)	----------	0.012 (1.5)	----------	0.025 (1.7)	**0.65	**0.59
(2) a. All industries	0.090	**0.037 (3.1)	0.0034 (1.0)	*0.0073 (1.7)	0.013 (1.6)	----------	----------	**0.78	**0.75
b. Motor vehicles excluded	0.097	**0.032 (2.7)	0.0010 (0.3)	*0.0080* (1.9)	*0.014 (1.8)	----------	----------	**0.65	**0.60
(3) a. All industries	0.053	**0.035 (2.9)	0.0093 (1.6)	----------	0.011 (1.3)	−0.00010 (0.3)	*0.028 (1.9)	**0.78	**0.75
b. Motor vehicles excluded	0.057	**0.031 (2.4)	0.0085 (1.5)	----------	0.011 (1.3)	−0.00016 (0.4)	*0.027 (1.8)	**0.65	**0.58

*Indicates coefficient is statistically significant at the 95-percent level.
**Indicates coefficient is statistically significant at the 99-percent level.

Figures in parentheses are t values.

In addition, if high profits lead to high advertising expenditures, we should expect that industries which have high profits for reasons other than product differentiation (e.g., concentration or technical entry barriers) would tend to have high advertising expenditures as well. Yet, as we noted above, advertising is only weakly correlated with the other dimensions of market structure.

CONCLUDING COMMENTS

On the basis of these empirical findings, it is evident that for industries where products are differentiable, investment in advertising is a highly profitable activity. Industries with high advertising outlays earn, on average at a profit rate which exceeds that of other industries by nearly four percentage points. This

differential represents a 50 percent increase in profit rates. It is likely, moreover, that much of this profit rate differential is accounted for by the entry barriers created by advertising expenditures and by the resulting achievement of market power.

We note also the significant joint impact on profit rates of concentration and the entry barriers created by scale economies and high capital requirements. Although the composite effect of these factors is clearly important, a more precise indication of the distinct effect of any of these variables is hazardous because of the high degree of collinearity. As would be expected, the rate of growth of demand has an important positive impact on profits. Models which incorporate these variables fit the underlying data reasonably well, accounting for approximatey 50 per cent of the variation in industry profit rates.

These empirical results suggest that factors which promote product differentiation may be as important as those which influence the size distribution of firms in terms of their effect upon the achievement of market power. Current policies which tend to emphasize the role played by concentration may need to be supplemented by those concerned directly with the nature and extent of product differentiation. Policies dealing with these matters would be an important component in a general policy designed to promote competition.

APPENDIX

DATA SOURCES AND TECHNICAL ADJUSTMENTS

The industry data used are reported at or aggregated to the level of I.R.S. "minor industries," which are roughly comparable to S.I.C. three-digit industry groups. The source for each variable is listed in table A1.

The sample was chosen originally to gain complete coverage of all consumer goods industries. All "miscellaneous" industries were eliminated, however, because of the obvious conceptual problems. In addition, three other industries were dropped from the sample: newspapers, while technically a manufacturing industry was considered to have sufficient "service" elements to make its inclusion inappropriate; petroleum refining, because of the unusual statistical problems which result from the tax treatment of mineral depletion; and motor vehicle parts, because of the lack of comparable Census data. Average profit rates and advertising-sales ratios for the remaining 41 industries are presented in table A2.

The variables are defined and explained in the text. The calculation of the technical entry barrier variables and the rate of growth of demand involved using both Census and I.R.S. data. The various specific adjustments made to reconcile data drawn from these two sources and reported at different levels of aggregation are described in the next two sections.

Technical Barriers to Entry

These variables are based on data from the 1954 *Census of Manufactures*. To carry out these computations, it is necessary to relate industries as defined by the Census Bureau to those of the Internal Revenue Service. This is done on the basis suggested by the Census Link Project.[41]

Within S.I.C. four-digit industries, average plant size among the largest plants which account for 50 per cent of industry output is used as the estimate of minimum efficient plant scale (MES). Data on shipments are used in all cases where available. In the few remaining cases, the calculations are based on value added. When the ratios of MES to industry output are obtained, the average percentage among component four-digit industries within the relevant I.R.S. industry is calculated, using shipments as weights where available and value added as weights elsewhere.

In determining the capital requirements variable, the scale of an efficient plant is measured in most instances by the value of the shipments but in a few by value added. In the latter cases, these figures are multiplied by the ratio of shipments to value added for the same four-digit industry but in a later year.

---

[41] Bureau of the Census, *Enterprise Statistics* (1958), Part 3.

TABLE A1.—SOURCES OF DATA

Variable	Source
(1) Profit rate	Internal Revenue Service Source Book of Statistics of Income. Average values for 1954–57.
(2) Advertising	Do.
(3) Concentration:	
(a) Trichotomous and dichotomous classifications.	Carl Kaysen and Donald F. Turner, Antitrust Policy, statistical appendix.
(b) Continuous 4 firm ratio	George J. Stigler, Capital and Rates of Return in Manufacturing Industries, 206–215.
(4) Economies of scale relative to market	1954 Census of Manufactures.
(5) Absolute capital requirements	1954 Census of Manufactures and Internal Revenue Service Source Book of Statistics of Income.
(6) Rate of growth of demand	Internal Revenue Service Source Book of Statistics of Income.
(7) Local market dummy variables	Carl Kaysen and Donald F. Turner, Antitrust Policy, statistical appendix.

When estimates of MES measured in shipments for all four-digit industries are obtained, these are averaged, using value added as weights, to derive the value in the larger I.R.S. industry. These averages are then multiplied by the appropriate assets-sales ratio for the I.R.S. industry, and the resulting figures used to represent the level of capital required for efficient entry.

In the case of the motor vehicle industry, Census data are unavailable. Bain's estimates for this industry are therefore used for both the extent of scale economies and the level of capital requirements.

For the three regional industries—soft drinks, dairy products, and bakery products—it is assumed that the appropriate market is the typical large metropolitan area. Output data are not available for four-digit industries by standard metropolitan area. Consequently, value added data for larger three-digit groupings are used. Average value added in large metropolitan areas for the three-digit industry [42] is multiplied by the ratio of total national shipments in the relevant four-digit industry to total value added in the associated three-digit industry to obtain the estimate of four-digit industry shipments within the typical local market. This figure is then used as the denominator in the estimate of M.E.S. to market for the four-digit industry.[43] Where necessary, the resulting estimates are aggregated to the I.R.S. minor industry level as described above.

Rate of Growth of Demand

This variable is the ratio of I.R.S. gross sales in 1957 to that in 1947. In a few cases, however, I.R.S. data are not available for both years and alternative procedures are used.

In 1947, the I.R.S. industries "Cigars" and "Cigarettes" are aggregated as "Tobacco Manufactures." To disaggregate the reported figure for gross sales, Census data on value of shipments (excise taxes excluded) for 1947 were examined. The ratio of shipments in "Cigars" to total tobacco manufacturing is multiplied by I.R.S. gross sales for tobacco manufactures, and the resulting product used to denote gross sales in "Cigars." Gross sales in "Cigarettes" is obtained residually.

[42] Output data were available for 56 standard metropolitan areas for the beverage industry (S.I.C. 208); 59 such areas for bakery products (S.I.C. 205); and 61 such areas for dairy products (S.I.C. 202).

[43] These corrections were made for the following subindustries: S.I.C. 2081—bottled soft drinks: S.I.C. 2021—creamery butter; S.I.C. 2027—fluid milk and other products; S.I.C. 2051—bread and related products.

TABLE A2.—AVERAGE PROFIT RATES AND ADVERTISING-SALES RATIOS IN 41 CONSUMER GOODS INDUSTRIES, 1954–57

[In percent]

	Profit rate	Advertising sales ratio
(1) Soft drink	10.0	6.2
(2) Malt liquors	7.2	6.8
(3) Wines	7.3	5.2
(4) Distilled liquors	5.0	2.1
(5) Meat	4.6	.6
(6) Dairy	7.9	2.2
(7) Canning	6.4	2.9
(8) Grain mill products	7.0	1.9
(9) Cereals	14.8	10.3
(10) Bakery products	9.3	2.9
(11) Sugar	5.8	.2
(12) Confectionary	10.6	3.5
(13) Cigars	5.3	2.6
(14) Cigarettes	11.5	4.8
(15) Knit goods	3.8	1.3
(16) Carpets	4.5	2.0
(17) Hats	1.6	2.2
(18) Men's clothing	5.9	1.2
(19) Women's colthing	6.1	1.8
(20) Millinery	−1.3	.8
(21) Furs	5.7	1.0
(22) Furniture	9.7	1.5
(23) Screens and venetian blinds	9.3	1.6
(24) Periodicals	11.7	.2
(25) Books	10.1	2.4
(26) Drugs	14.0	9.9
(27) Soaps	11.7	9.2
(28) Paints	9.9	1.5
(29) Perfumes	13.5	15.3
(30) Tires and tubes	10.2	1.4
(31) Footwear	7.6	1.5
(32) Handtools	11.4	4.2
(33) Household and service machinery (not electrical)	7.3	1.9
(34) Electrical appliances	10.3	3.5
(35) Radio, TV, and phonograph	8.8	2.2
(36) Motorcycles and bicycles	5.2	1.1
(37) Motor vehicles	15.5	.6
(38) Instruments	12.0	2.0
(39) Clocks and watches	1.9	5.6
(40) Jewelry (precious metal)	5.3	3.2
(41) Costume jewelry	1.4	4.0

I.R.S. data for "Screens and Venetian Blinds," in 1947, are included in "Miscellaneous Furniture." As a result, Census value added data for both 1947 and 1957 are used. Similar data problems exist in the case of "Drugs," "Perfumes," "Instruments," and "Costume Jewelry." In all of these instances Census data in 1947 and 1957 are used. With regard to "Perfumes" and "Costume Jewelry," information on value of shipments is available in both years and is therefore used, while value added data are used in the other cases.

Concentration

The classification of industries into Kaysen and Turner concentration classes is based on data for 1954, and their definitions are used in most cases. Where I.R.S. minor industries include a number of Kaysen and Turner markets, weighted averages based on value of shipments are used to determine the appropriate classifications.

Further adjustments are made to account for the local market character of three industries: soft drinks, dairy products, and bakery products. In addition, the soap industry falls just below the boundary between Type I and Type II Oligopolies. Since the same major firms are dominant in the important product lines, the industry is reclassified as a Type I Oligopoly. The radio and television industry falls just above the same boundary. It is classified as a Type II Oligopoly since the largest single subindustry falls into that category. For appliances, the Kaysen and Turner definition differed substantially from that used by the Internal Revenue Service. After an appropriate adjustment the concentration data indicated that this industry is a Type II Oligopoly.

The classification of industries according to whether the eight-firm concentration ratio exceeded or fell short of the 70 per cent level is based on the concentration ratios presented by Kaysen and Turner.

(Whereupon, at 11:55 a.m., the hearing was adjourned subject to call of the Chair.)

81-280—68—pt. 5——36

APPENDIXES

APPENDIX I—CORRESPONDENCE BETWEEN MR. BENJAMIN GORDON AND DR. P. M. COSTELLO RE TECHNOLOGICAL PROGRESS IN THE PHARMACEUTICAL INDUSTRY

MARCH 6, 1968.

Dr. P. M. COSTELLO,
Department of Economics, Smith College,
Northampton, Mass.

DEAR DR. COSTELLO: It is my understanding that you have done considerable research concerning technological progress in the pharmaceutical industry.

I should be very grateful if you would submit to our Monopoly Subcommittee a brief summary of your findings in this field. Your contribution on this aspect of the subject would be very useful in our study of this industry.

Your cooperation is appreciated.

Sincerely,

BENJAMIN GORDON,
Staff Economist.

SMITH COLLEGE,
Northampton, Mass., March 13, 1968.

Mr. BENJAMIN GORDON,
Staff Economist, Select Committee on Small Business,
U.S. Senate, Washington, D.C.

DEAR MR. GORDON: In reply to your letter of the 6th, I am enclosing a paper which examines the pace of technological progress in the ethical drug industry. This study was developed from research for my doctoral dissertation which considered the narrower question of patents in the antibiotic segment of the industry.

Sincerely,

P. M. COSTELLO.

TECHNOLOGICAL PROGRESS IN THE ETHICAL DRUG INDUSTRY*

(By P. M. Costello, Assistant Professor of Economics, Smith College, Northampton, Mass.)

One critical test of an industry's market performance is its rate of technological progress. For the ethical drug industry this dimension of performance is of added significance in view of the increasing role played by drugs in the management of human diseases. This paper examines a sample of 528 new drugs innovated between 1945 and 1965. The first section presents estimates, based upon medical evaluation and opinion, of the proportion of these innovations that represent significant advancement in the art of chemotherapy. The second section considers the problems of constructing an adequate criterion to judge the pace of advancement from the antibiotic segment of the industry.

I. THE RECORD OF TECHNOLOGICAL PROGRESS

The sample of new drugs introduced to the market between 1945 and 1965 was compiled mainly from the American Medical Association's New and Nonofficial Drugs. In general the drugs selected for analysis were new chemical structures excluding combinations of old drugs and new dosage forms. The exclusion of the latter two categories was founded on the assumed lower order of scientific ability required to produce such products. But even within the category of new chemical structures major differences in the effectiveness of these drugs were immediately apparent.

---

*An earlier draft of this article received valuable comment from Professors Joel B. Dirlam and Robert Britt.

In an attempt to meet the qualitative problem two categories are defined. The first is designed to include drugs which were in one degree or another an improvement over older products. These varied from drugs which were considered medical breakthroughs to drugs which reduced the incidence of undesirable side effects observed with established chemotherapeutic agents. Drugs which fell in this category were counted as examples of technological progress.

The second category includes those drugs which from a technical point of view were no advancement over older products. The range in this category is from drugs which had the same degree of effectiveness and the incidence of side effects as older products to drugs which were removed from the market or limited in their application because of undesirable side effects. Also included in this category are drugs where medical opinion was divided or clinical experience too limited to form a statistically valid sample of the value of the drug. Drugs falling in this category were counted as examples of product differentiation which exhibit no significant technological progress.

To classify each of the drugs in the sample the Medical Letter[1] and the AMA's New and Nonofficial Drugs were selected as authoritative publications that best summarized clinical experience with each drug. Thus the classification of a drug is directly dependent upon its technical efficiency relative to a particular medical problem. This scheme gives no weight to sales volume as an index of the relative value of a drug. This recognizes that physicians are responsive to the promotional claims of the manufacturers and will prescribe a new drug before sufficient clinical experience is amassed to establish the comparative value of the new drug relative to older products. As a result a drug significant in the management of a disease of low incidence in the population and a correspondingly low volume of sales receives the same weight as a drug which cures diseases of high incidence and high volume of sales.[2]

Biases in the sample are undoubtedly numerous and difficult to evaluate and could operate in either direction. First, the source publications do not clearly indicate the basis of selecting a particular drug for discussion. Inclusions may well be in response to the promotional claims of the manufacturer with the recognition that the majority of practicing physicians have neither the facilities, the time, the training, nor the range of patients necessary to conduct statistically valid tests of such claims. Second, definitive judgment of the place of a particular drug in the physician's tool kit may require several years of experience before the population treated is sufficiently large to expose undesirable side effects. Thus, a drug currently considered significant may, with further testing, be found to have only a restricted application. Third, drugs now considered insignificant may, through modifications in techniques of administration, or in combinations with other drugs, be found valuable in the management of specific diseases. Finally, there is a bias in restricting the sample to new chemical structures. It is possible that combinations of old drugs, for example, may increase the number of diseases controllable through chemotherapy. But there is no evidence to date that this technique has produced anything of importance.

In Table I of the sample of 528 new chemical drugs is divided according to the classification scheme outlined above. Column I contains estimates of the total number of new chemical structures innovated yearly between 1945 and 1965.

---

[1] The Medical Letter is a publication designed to acquaint the practicing physician with the latest information on drug products.

[2] For a different approach to the qualitative problem see W. S. Comanor, "Research and Competitive Product Differentiation in the Pharmaceutical Industry in the United States," *Economics*, Vol. 31, November 1964, pp. 372–84; and "Research and Technical Change in the Pharmaceutical Industry," *Review of Economics and Statistics*, Vol. 47, May 1965, pp. 182–90. Comanor's investigation of the whole range of new drug innovations weights each innovation by its sales volume for the first two years following its introduction. My objection to following this approach in the present study is that the time required to develop experience with any particular new drug to support a definitive opinion on its technical worth seems to exceed the two-year period following introduction.

TABLE I.—ESTIMATE OF TECHNOLOGICAL PROGRESS FOR NEW CHEMICAL DRUGS, 1945–65

Year	Total new chemical drugs (I)	Estimate of the number representing product differentiation (II)	Estimate of the number representing technological progress (III)	Col. III as a percent of col. I (IV)
1945	3	2	1	0.33
1946	16	14	2	.13
1947	16	11	5	.31
1948	14	10	4	.29
1949	28	21	7	.25
1950	25	23	2	.08
1951	21	17	4	.19
1952	22	21	1	.05
1953	38	33	5	.13
1954	35	33	2	.06
1955	27	24	3	.11
1956	27	26	1	.04
1957	38	35	3	.08
1958	41	37	4	.10
1959	52	46	6	.12
1960	34	28	6	.18
1961	18	16	2	.11
1962	20	19	1	.05
1963	16	14	2	.13
1964	16	15	1	.06
1965	21	20	1	.05
Total	528	465	63	.12

Industry figures for new chemical drugs are available for the period 1951–61.[3] Comparing these figures with those in Column I indicates that the industry data exceeds the sample data by a yearly average of 10.8 drugs. In part this discrepancy is due to differences in the year of introduction reported for certain drugs and perhaps due in part to differences in the definition of new chemical drugs used by the industry and the AMA. It seems likely however, that if any of these innovations reported by the industry were important they would have been reported to the medical profession through the AMA's publications. The effect of including the industry data here would be to reduce the proportion of total innovation that represented technological advancement.

Beginning in the late 1940's there is a persistent increase in the number of drugs marketed each year until 1959, followed by a marked decline up through 1964. The series representing technological progress, Column III, however fails to increase appreciably. For the period 1945–50 the average ratio of innovations to advances is 5 to 1, while for the period after 1950 this ratio increases to an average of 10 to 1. Both series exhibit a decided break appearing in 1960 and 1961 respectively.

These broad movements can in part be explained by a combination of structural factors internal to the industry and external institutional factors. For the 1945–50 period the industry engaged in a process of forward vertical integration. The objective of this integration was better control over prices through the elimination of the firms solely engaged in packaging. The major example was set by Parke Davis and Lederle, the innovators of the first broad spectrum drugs. Both of these firms were fully integrated into the retail pharmacy and hospital markets and held strong patents on their innovations. In refusing domestic licenses on these products the traditional marketing pattern in the industry was altered. With the elimination of the packagers, direct control over price became possible. This example of the use of the patent and forward integration was undoubtedly a factor in Merck's merger with Sharp and Dohne as well as Pfizer's entry by internal expansion.

The structural change in the industry suppressed one form of competition—imitation from packagers or from integrated firms under license agreements. The alternative avenue open was to invent around the patent to produce a close substitute. Thus the increase in the proportion of innovations which represented product differentiation was partly in response to the closing of one method of competition.

[3] Pharmaceutical Manufacturers Association, *Prescripting Drug Industry Fact Book* (Washington, PMA), 1963, pp. 2–5.

The decline in both innovation and technological progress observed in the early 1960's can be explained as the industry's reaction to criticism generated by the Kefauver investigation and the stricter controls imposed by the FDA during this period. Table II indicates that the rate of increase in expenditures for research and development fell measurably in 1960, leading to a decline in innovations as well as in the number that represent technological progress.

TABLE II.—RESEARCH AND DEVELOPMENT EXPENDITURES AND EMPLOYMENT FOR ETHICAL DRUGS

Year	Expenditures (millions)	Percentage change	R. & D. employment (thousands)	Percentage change
1956	$94			
1957	104	0.11	4.7	
1958	128	.23	5.1	0.08
1959	154	.20	5.9	1.6
1960	163	.06	6.0	0.2
1961	181	.11	6.2	0.3
1962	198	.09	6.7	0.8
1963	215	.09	6.8	.01
1964	238	.06	7.2	.06
1965	268	.13	7.6	.06

Source: National Science Foundation, Review of Data on Science Resources, Washington, various issues.

In general, if technological progress is measured as the percentage of innovations which find general medical acceptance, it is clear that over time the proportion of drugs that can be classified as representing meaningful progress declines. Or to state the case somewhat differently, the series on the number of new drugs of lasting value appears rather constant over time in spite of the marked increase in the number of innovations. These data then tend to support the observation that while the industry has increased its expenditures on research and development, much of the added effort appears to have resulted in product differentiation.

II. CRITERION FOR JUDGING PERFORMANCE

Attempts to judge the "fairness" of prices or profits are generally made on the basis of comparisons of price to costs of production, or profits of one firm in comparison with those of another. But with technological progress as the performance variable no such objective criterion exists. In the present case we do not know whether the proportion of innovations representing technological progress would have been higher or lower if the industry did not, for example, have access to patent protection. If the research resources which were directed to patent circumvention had been used to attack more significant problems, it is possible that the research output would have been higher. But it is also possible that given the quality of these resources, product modification was the best they were capable of in their employment in the industry. The problem is of course a familiar one, in that we cannot rerun the industry's reactions under different structural characteristics.

An alternative approach is to compare progress under different market structures. But even here competing hypotheses of the effect of structure on performance are in existence and debated. The Schumpeter version that only firms with some degree of market power have the resources to innovate is opposed by the hypothesis that innovation offers competitive firms escape from the rigors of competition. The antibiotic segment of the drug industry does shed some light on these competing hypotheses and can aid in the judgment of whether progress has been adequate.

Table III presents data on the innovation record for the antibiotic segment of the industry. Total innovations are divided between the so-called narrow and broad spectrum drugs corresponding to a competitive segment and a monopolized segment of the market. In the narrow spectrum market approximately 30 percent of these innovations represented some degree of technological progress, while in the broad spectrum market one-half of the innovations were significant in this respect.[4] In light of different market structures associated with these

[4] The four broad spectrum drugs and their classification are as follows: Both chloramphenicol and chlortetracycline are counted as evidence of technological progress. Oxytetracycline and tetracycline are counted as evidence of product differentiation as both are very close substitutes for chlortetracycline.

two markets, the differences in the proportion of innovations representing an advance lends initial support to the hypothesis that monopoly is a necessary condition for progress. But examination of the development of the broad spectrum market and its operation weakens this conclusion.

TABLE III.—PRODUCT INNOVATION IN THE NARROW AND BROAD SPECTRUM ANTIBIOTIC MARKETS

Year	Narrow spectrum innovations	Broad spectrum innovations
1945	1	
1946	1	
1947	2	
1948	1	1
1949	1	1
1950	3	1
1951	1	
1952	1	
1953	2	1
1954	2	
1955	3	
1956		
1957	3	
1958	3	
1959		
1960	3	
1961		
1962	2	
1963	1	
1964	2	
1965	2	
Total	34	4

The origins of the antibiotics are found in Alexander Fleming's discovery of penicillin in 1928 and in the development research undertaken by a group headed by H. W. Florey in England in the late 1930's. England's involvement in the war and problems of mass producing penicillin caused Florey to seek commercial sources of supply in the United States. The majority of firms contacted by Florey expressed little interest in pursuing the research necessary for large scale production. But he was successful in his contact with the Department of Agriculture's Northern Regional Research Laboratory (NRRL). With the entry of the United States into the war a cooperative program was instituted by the Committee on Medical Research (CMR). The program included a number of drug firms, university and government laboratories, as well as laboratories in England. The commercial laboratories of the drug firms initially concentrated their research on the synthetic production of penicillin,[5] while the noncommercial laboratories concentrated on the problem of mass production. The development of improved mold stains by the university laboratories and the 'deep fermentation' production process by the NRRL essentially solved the key problems of mass production.[6]

By the mid-1943 the CMR shifted the emphasis in the penicillin program from research to production. To overcome the reluctance of the firms to abandon their synthetic research and adopt the available technology, the program was expanded to include firms not initially under contract with the government

[5] It has been argued that the drug firms concentrated on the synthetc approach at the expense of the production technology suggested by the NRRL. The reason offered for this approach was that product patents were unavailable on penicillin, and the Department of Agriculture held the key process patents. Richard Harris, *Annals of Legislation, the Real Voice,* The New Yorker, for March 14th, 21st and 28th, 1964. Especially page 69 of the March 14th issue, ". . . the firms were too busy trying to corner patents on the various processes in the production of penicillin to produce much of it, and the government began to press them to work together." A similar conclusion was reached in Federal Trade Commission, *Economic Report on Antibiotic Manufacture* (Washington, Govt. Printing Office, 1958), pp. 37–38. In defense of the firms involved it should be noted that historically, the synthetic production of drugs had made the 'natural technology' obsolete. For a description of Merck's experience and others see, Tom Mahoney, *The Merchants of Life* (New York, Harpers, 1959), Chaps. 11, 13, 16.

[6] Until the mid-1940's penicillin was produced by two methods. The less efficient 'surface technique' involved the growth of penicillin on the surface of the culture medium. With the 'deep fermentation technique' the medium was aerated allowing growth throughout the medium.

through the CMR and pressure on those under contract to begin production. In addition, government funds were provided for plant construction and modification,[7] and as a result, the War Production Board permitted the sale of penicillin through normal trade channels in March of 1945.

After the removal of government restriction on the safe of penicillin, market forces operated to reduce prices. Tariff Commission data indicate that between 1945 and 1950 the realized price of penicillin fell from $3,955 to $282 per pound, while production increasing from 11,746 to 329,746 pounds.[8]

The behavior of penicillin prices can be explained by two factors—conditions of entry and the existence of a bulk market. A product patent on penicillin was unavailable, and the process patents were held by the Department of Agriculture, which followed a policy of freely licensing applicants; entry could not be restricted in manufacturing or selling based upon patents. Available evidence indicates that there were 16 penicillin suppliers in 1945, and 13 in 1950.[9]

The second factor was that several of the major drug producers were not integrated forward into direct selling to hospitals and retail pharmacies. These firms generally sold a significant proportion of their output through the bulk market to packagers who resold under their own brands. The FTC study reported 27 packagers in 1950, increasing to 45 by 1956, concluding that entry at this level was relatively easy.[10] Thus packagers purchased their supplies in the open market and sold in competition with the integrated producers.

The emphasis on price competition in the penicillin market was not at the expense of product improvement. Various modifications of penicillin were introduced. For example, Squibb introduced the first orally effective penicillin in 1945, Wyeth introduced benzathine penicillin which has a slower rate of absorption in 1951, and in 1955 Lilly marketed penicillin V claiming higher potency. Product patents were generally obtained on these specialties but usually only after interference proceedings had been declared by the Patent Office. Settlement took the form of withdrawal of a competitive patent application in return for a license to produce and sell under any patents that be issued. While data on realized prices of these products are not reported separately by the Tariff Commission, these prices generally drifted downward following the price of the basic drug although at a somewhat higher level.

Thus in the penicillin market the underlying condition of freedom of entry and the bulk market produced a market competitive in price and with product improvement.

The second major antibiotic following penicillin was streptomycin, discovered by Selman Waksman at Rutgers University in 1943. Under the terms of the research contract between Waksman and Merck, the latter received exclusive commercial rights to this drug. To exploit its position, however, Merck had to first convince the Patent Office that streptomycin was sufficiently important to warrant patent protection,[11] and second, to integrate forward into the retail market to avoid dependence upon the bulk market. Both steps were necessary. For with the sale of a patent product to packagers, the patent monopoly was exhausted and price cutting could legally begin.[12]

The demonstration of the potential importance of streptomycin resulted in Waksman's request to Merck that right of commercial exploitation be given up. The sequent agreements provided for the assignment of the patent application to a nonprofit foundation at Rutgers, and royalty abatement of $500,000 to cover Merck's development costs. In 1948 the patent was issued to the foundation which followed an unrestrictive licensing policy.[13] Seven firms in addition to Merck were licensed to manufacture streptomycin. Four of these firms, including Merck,[14] were not integrated forward, selling primarily in the bulk market to

[7] FTC, *op. cit.*, pp. 47–56.

[8] U.S. Tariff Commission. *Synthetic Organic Chemicals, United Production and Sale,* Reports No. 159 and 167 (Washington. Govt. Printing Office), various issues.

[9] The decline in the number of manufacturers was largely due to their adoption of inefficient production techniques. With the end of military purchases and the onset of price competition these firms would not cover costs and transferred their resources to other markets.

[10] FTC. *op. cit.*, p. 66.

[11] Prior to streptomycin the Patent Office took the position that the isolation of products of nature did not rise to the level of invenion and hence were unpatentabble. Merck successfully argued that antibiotic products were transitory in nature and that their isolation constituted invention within the purpose of the statutes. F. Cacciapaglia, Jr., and H. B. Reckman. *The Proposed Drug Industry Antitrust Act—Patents, Pricing and the Public,* The George Washington Law Review, Vol. 30, June 1962, p. 890.

[12] The doctrine that the first sale of a patented product exhausted the patent monopoly was set out in Boston Stores Co. v. American Gramophone, 246 US 8 (1917).

[13] FTC. *op. cit.*, p. 229.

[14] Merck did not enter the retail market until its 1952 merger with Sharp and Dohne.

packagers. The market results were similar to those in the case of penicillin. Tariff Commission data indicate that the realized price of streptomycin fell from $2,866 per pound in 1946 to $160 in 1950. Competitive market conditions did not prevent Merck from further research on streptomycin. In 1948 dihydrostreptomysin was marketed as an improved version of streptomycin, and the product patent was issued to Merck in 1950. Six firms were licensed, three of which were bulk sellers. The price of dihydrostreptomycin has closely followed that of streptomycin.

The marketing of the broad spectrum antibiotics in the late 1940's and early 1950's ended the competitive characteristic of the postwar antibiotic market. In late 1948 Lederle Laboratories, a division of American Cyanamid, introduced the first broad spectrum antibiotic, chlortetracycline, closely followed by Parke Davis with chloramphenicol. As both firms were fully integrated into direct selling the bulk market was bypassed. With the issuance of the product patents competition could be legally excluded. Evidence through 1960 indicated that neither company made bulk sales to packagers and all requests for domestic licenses were refused. The prices of these drugs were reduced from their introductory level of $15 per bottle of 16-250-milligram tablets, reaching a floor in 1951 at $5.10 and remaining constant until the opening of the Kefauver investigation in 1960.

In 1950 the Chas. Pfizer Company introduced the third broad spectrum drug oxytetracycline. While waiting for FDA clearance a sales force was organized, thereby avoiding sales in the bulk market. With the issuance of the product patent all requests for licenses were rejected. The marketing of oxytetracycline caused Lederle to reduce the price of its entry, with Parke Davis meeting the price cut. The last major broad spectrum drug was tetracycline, introduced in 1953 by Lederle, but also produced by Pfizer and Bristol Laboratories and sold by Squibb and Upjohn as a result of patent interference claims. The introduction of tetracycline did not disturb the broad spectrum price level nor were domestic licenses granted to other producers or sellers.

Thus an examination of the early narrow spectrum market indicates a competitive market structure at both the manufacturing and packaging levels. Competitive behavior can be observed in price as well as in product innovation. The availability of patents and forward vertical integration provided a means by which the broad spectrum innovators could escape the rigors of competition.

The examination of the antibiotic segment of the industry points up quite clearly the major flaw in the measurement of technological progress employed earlier. If the market is subject to some degree or form of monopoly control where firms can reach agreement not to compete on the basis of price, and prices can be maintained at relatively high levels, there is a tendency not to disturb these agreements through product innovations. This has the effect of decreasing the base and overstating the proportion of innovations which constitute technological progress. Having achieved control over price and profits the major stimulant to further research is removed. There is evidence from antibiotics that once market control is established further research in the therapeutic area ceases. First, it seems doubtful that tetracycline, introduced in 1953, is the ultimate broad spectrum drug. The strong patent positions of the manufacturers and Pfizer's aggressive defense of its patent on tetracycline undoubtedly reduce the potential profitability of this segment of the market for firms contemplating entry through research. The broad claims allowed in antibiotic patents would be a major factor in discouraging entry directed research. The potential entrant to be free of patent infringement suits would be forced to discover antibiotic producing microorganism unclaimed by the established manufacturers. Surmounting this problem there remains the threat of infringement suits designed to harass. And unless the innovating firm is of equal size with its competitor the evidence indicates that a policy of harassment will be successful in forcing him from the market. Here one can cite the fact that Pfizer drove 33 smaller competitors from the tetracycline market between 1960 and 1965 by this method.

Second, there is evidence that Lederle, after the discovery and innovation of chlortetracycline abandoned research in this area until faced with serious competition from Pfizer in late 1952. In 1948 Lederle carried out experiments with chlortetracycline that produced a substance with antibiotic properties. In the patent dispute with Pfizer, Lederle claims that this substance was in fact tetracycline. But in 1948, Lederle dominated the broad spectrum market with chlortetracycline and evidently did not feel compelled to complete its experiment. But in 1952, with Pfizer's discovery of tetracycline Lederle ". . . resumed the 1948

work, and in December 1952 or January 1953 . . . produced tetracycline . . ." [15] This lends support to the argument that once a dominant position is established and protected by patents, interest in further innovation declines until the basis of the dominance is threatened.

Third, Parke Davis, one of the original innovators in the broad spectrum market, has never made any additional contributions.

The observations from the narrow spectrum market suggests that where competition remains an active force innovation continues. And in comparing the approximate 30 percent of these innovations which were found to be significant with the 12 percent for the industry as a whole indicates that competition is a major determinate of technological progress. In the broad spectrum market where competitive forces were suppressed one finds a high proportion of the innovations significant. But to conclude from this observation that monopoly promotes progress would be in error. First, the number of innovations were considerably fewer than in the narrow spectrum market. This in turn suggests that the agreement not to compete on price included an agreement not to compete in terms of product innovation as well. Second, an examination of the discovery of the broad spectrum drugs indicates that associated with the establishment of monopoly is a decline in research within the therapeutic area. And only when the basis of the monopoly is threatened by competition does research resume.

APPENDIX II—PAPER, "RISK AND CORPORATE RATES OF RETURN," BY I. N. FISHER AND G. R. HALL,[1] THE RAND CORPORATION, SANTA MONICA, CALIF.

(Cited to text, p. 1820, supra.)

I. INTRODUCTION

Although economists have great interest in the correlation between risk and profits, few studies have attempted to quantify the relationship.[2] Consequently, this paper considers the concept of risk differentials in corporate profit and proposes a model for measuring them. Using this model, the risk-rate of return relationship was estimated for a sample of firms in various industry groups. For each industry group, average risk-adjusted rates of return were also obtained.[3]

Risk is defined as the inability to predict the outcome of a forthcoming event with complete certainty. Entrepreneurs are viewed as making decisions in the face of uncertaintyly on the basis of probabilistic expectations about future outcomes.[4] If certaintly is a situation where the entrepreneur's anticipation will assuredly be fulfilled, then uncertainty can be measured by the likelihood that the actual outcome will differ from the anticipated outcome.

The foregoing definition accords with economic models of risk (Refs. 6, 23, 24), and it suggests studying risk by examining distributions of corporate rates of return. Specifically, this approach intimates that the risk-rate of return relationship can be analyzed statistically in terms of the relationship between the mean rate of return and higher moments of the distribution.[5]

---

[15] American Cyanamid Company, *Proposed Findings of Fact,* Before the Federal Trade Commission, Docket No. 7211, 1960. Vol. I, p. 16.

[1] Any views expressed in this paper are those of the authors. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This paper was prepared for presentation at the 1967 Annual Meeting of the Econometric Society, December 27–30, Washington, D.C. Helpful comments and suggestions were provided by F. D. Arditti, E. A. Greenberg, R. E. Johnson, L. J. Kleiger, M. H. Kosters, A. Phillips and I. H. Plotkin.

[2] Examples of the wide and diverse theoretical literature on risk and profits are Refs. 6, 8, 11, 14, 17, 18, 20, 27, 32, and 34. Two important empirical studies of risk and rates of return are Refs. 4 and 30. In this paper "risk" will not be distinguished from "uncertainty." Compare Refs. 4, 6, 9, 11, 18, and 34.

[3] The 88 firms included in the sample were selected from *Fortune* magazine's list of the 500 largest industrial firms. Eleven industry groups were selected, comprising a variety of different types of industries. Firms in each group were chosen so that both middle-sized, as well as large, firms would be included.

[4] Economic theory contains two approaches to this problem : in one approach, the decision-maker balances the various moments of the probability distribution of potential outcomes on the basis of his utility function (Refs. 21 and 28), while in the other, the decision-maker chooses among a set of dated financial claims defined over all future states-of-the-world (Refs. 13, 14). We adopted the first, since the data do not justify the more elegant approach.

[5] It is assumed that the present and future sets of profit-generating opportunities for each firm are determined exogenously. That is, a firm may select opportunities but cannot influence the composition of any set of potential investments. Without this assumption, the concept of risk becomes more complex. If firms can influence the investment-choice set, however, presumably the observed variance of profits would decrease over time. Thus, concentration on fairly long periods of time and a large sample of firms should lessen the likelihood of this possible effect biasing the statistical results.

II. THE MODEL

Assume that firms maximize not profits, but expected utility, and let U(P+W) be the firm's utility function.[6] Utility is a function only of earnings, P (a random variable), and net worth, W. The risk premium, R(P,W), is that amount required to make the entrepreneur indifferent between the expected value of the uncertain earnings, E(P+W), and the certain amount E(P+W)—R(P.W), corresponding to the expected utility of the uncertain earnings (Refs. 8, 26).

Earnings distributions and utility functions are not important per se; it is their interaction that determines the risk component of profits. Suppose that both the probability distribution of potential earnings and the firm's utility function are known (illustrated in Fig 1(b) for a risk-averse firm).[7] Assume the probablity distribution is curve (1). Both the probability distribution of utilty, shown as (1) in (a), and its expected value, $E(U_1)$ are easily derived. Note that, although the probability distribution of earnings is symmetric about the expected value, E(P), the distribution of utilities is skewed to the left. This occurs because the utility function is concave, resulting in a non-linear transformation from earnings into utility. The expected value of the utility distribution, $E(U_1)$, is less than the utility of expected earnings, U(E(P+W)), and the difference, translated into monetary terms, is the risk premium E(P)—P*.

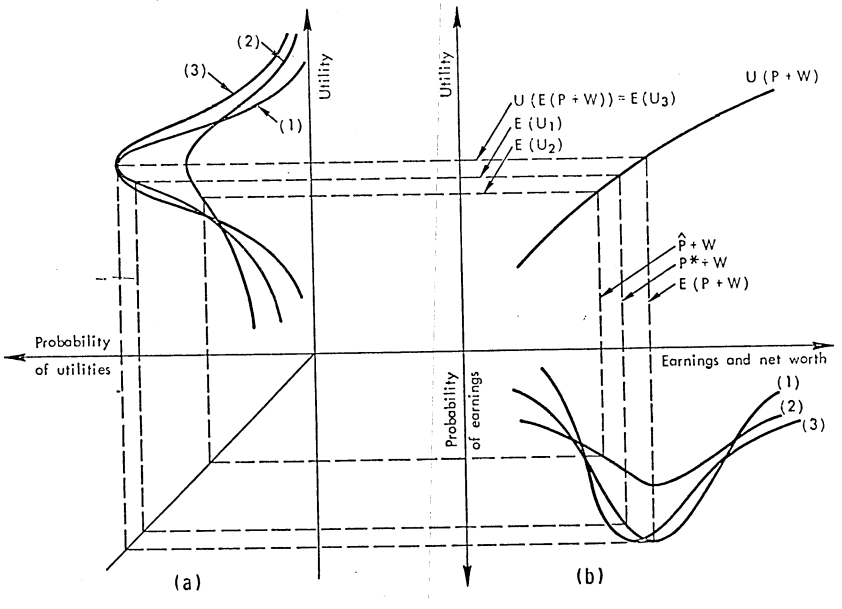


Fig. 1—Effect of dispersion and skewness on risk premium

Now suppose that the probability distribution is not curve (1) but curve (2). This distribution is also symmetric about the mean, but the variance is larger. The distribution of utilities is curve (2) in (a) and, as before, it is not symmetric

[6] Here P refers to profits in the sense of increments to net worth, rather than the profit rate.
The question of "whose" utility function is moot. There are various candidates, e.g., managers, stockholders, the chief executive officer, as well as others (Ref. 10). It is assumed that each entrepreneur (management) is interested in maximizing the expected utility of the net worth of the firm on the basis of his judgments about stockholders' preferences. Such an assumption permits us to explore the relationship of uncertainty to earnings without having to deal with the complexities of adding stockholders' utility functions, or the Modigliani-Miller view that stockholders can lever portfolios to offset corporate management decisions about risk (Refs. 22, 24. See also Ref. 6).

[7] If the firm is averse toward risk, the utility function is concave. This requires that $U' > 0$ and $U'' < 0$, or that utility increase with earnings and net worth, but at a decreasing rate.

about its mean, $E(U_2)$. The important point, however, is that the expected utility has decreased as a result of the increased dispersion of the earnings distribution. As a result, the risk premium, $(E(P)-P) > (E(P)-P^*)$. Consequently, greater variance in the distribution of earnings implies greater risk and, for risk-averse firms, leads to larger risk premiums. This implies that earnings should be larger, on average, for firms with greater variation in their earnings than for firms with little earnings variability.

Skewness may also have an important effect on the risk premium (Refs. 3, 12, 31). The entrepreneur may prefer positively-skewed earnings distributions because the likelihood of extremely low earnings is smaller. This, also, is illustrated in Fig. 1. Curve (3) in (b) has the same expected value as (1) and (2), but is skewed to the right. This function has been constructed so that the resulting distribution of utilities is symmetric about its expected value. In this example, skewness offsets variance and the risk premium is zero, i.e., $E(U^3) = U(E(P+W))$. Thus, positive skewness results in smaller risk exposure, while negative skewness leads to greater risk exposure, implying that earnings should be smaller, on average, for firms with earnings distributions positively skewed but larger, on average, for firms with negatively-skewed distributions.

The results of Fig. 1 suggest that once the form of the utility function is specified, risk exposure can be measured by characteristics of the probability distribution of earnings. The required risk premium becomes larger as the spread of the earnings distribution increases, but the premium decreases as the distribution becomes positively skewed.[8] This illustrates that risk exposure, as defined here, can be measured by characteristics of the firm's earnings distribution.

Before testing this hypothesis, one link in the discussion of the relationship between risk and earnings remains to be completed—that of the mechanism by which entrepreneurial preferences for risk and profits are translated into industry profit differentials or risk premiums and discounts.

Conventional economic theory indicates that with well-functioning capital markets the equilibrium rate of return will be identical among all activities. Entrepreneurs theoretically seek those investments yielding the largest rates of return. As capital is withdrawn from less profitable activities, the rates of return in such activities rise. Similarly, the inflow of capital into higher-yield investments forces the rates of return in these activities downward. Equilibrium occurs when the rates of return of investment are identical among all activities.

When risk is considered, the adjustment process is more complex. Because differences in risk exposure exist among alternative investments, entrepreneurs balance risk against expected rates of return. Capital, therefore, is transferred from low-return, high risk activities to high-return, low risk investments until an equilibrium, characterized by a set of risk premiums reflecting differences in risk exposure, is achieved. In this equilibrium, risk-compensated rates of return are equal among alternative investments, but observed or actual rates of return will differ by the amount of the risk premiums.

---

[8] This can be demonstrated formally in the following manner: Expand $U(P+W)$ in a Taylor series about the point $(\hat{P}+W)=E(P+W)$,

$$U(P+W)=U(\hat{P}+W)+U'(\hat{P}+W)(P-\hat{P})+\frac{U''}{2!}(\hat{P}+W)(P-\hat{P})+\frac{U''}{3!}(\hat{P}+W)(P-\hat{P}+\ldots.$$

Taking expected values and holding W, $\hat{P}$ constant,

$$E(U(P+W))=U(\hat{P}+W)+\sigma_p^2\frac{U''}{2!}(\hat{P}+W)+\sigma_p^3\frac{U''}{3!}(\hat{P}+W)+\ldots.$$

Rearranging terms, the difference between expected utility and utility of expected earnings is

$$U(P+W)-E(U(\hat{P}+W))=-\left(\sigma_p^2\frac{U''}{2!}(\hat{P}+W)+\sigma_p^3\frac{U''}{3!}(\hat{P}+W)+\ldots.\right).$$

Equation (3) is the risk premium, $R(P,W)$, and it becomes apparent that the second, third, and higher moments may affect the magnitude of the risk premium. Since $U''<0$ for a concave utility function, the risk premium must increase with larger variances. (The appropriate revisions for risk neutrality or risk preference should be apparent.) It is not, however, clear whether $U'''\gtrless 0$. If we assume that firms enjoy positive skewness (longshots), $U'''>0$ and the risk premium becomes smaller as skewness increases. Higher moments add little information about the characteristics of the distribution and are ignored. (See references 3, 26).

In short, we posit that capital markets respond to risk as they respond to expected rates of return. We should, therefore, expect to find a structure of risk-compensated rates of return that motivate or discourage investment. Part of the earnings differentials observed among alternative investments can be attributed to risk; these are the risk premiums that compensate for differences in risk exposure.

III. EMPIRICAL RESULTS

To test the hypothesis that profits are larger for firms with greater risk exposure, it is necessary to translate the theoretical definition of risk into statistical terms.[9] We can do this by assuming managers' anticipations, on average, are correct, thereby permitting the observed mean rate of return to be used as a proxy (Ref. 4). Risk exposure, as defined here, can then be measured by moments of the distribution of earnings.[10]

The risk variables were calculated from

$$\sigma_i=\left[\sum_{t=1}^{n}\frac{(r_{it}-\hat{r}_{it})^2}{n}\right]^{1/2} \tag{1}$$

and

$$S_i=\sum_{t=1}^{n}\frac{(r_{it}-\hat{r}_{it})^3}{n\sigma_i^3} \tag{2}$$

where

r_{it}=observed rate of return for firm i in year t;
$\hat{r}_{it}$=predicted rate of return from trend for firm i, year t;
σ_i=standard deviation of rates of return about trend, firm i;
S_i=skewness about trend for firm i;

and n is the number of years included in the sample.

The model can now be stated explicitly as

$$\bar{r}_i=r_0+b_1\sigma_i+b_2S_i \tag{3}$$

where

$\bar{r}_i$=average rate of return on net worth for firm i;
r_0=intercept; and b_1, b_2 are the coefficients of the standard deviation and skewness, respectively—the risk coefficients.

The signs of these coefficients are expected to be

$$b_1>0$$

$$b_2<0.$$

Estimates of the relationship between average rate of return and risk exposure appear in Table 1. Regressions (1) and (2) show the individual contribution of standard deviation and skewness in explaining variations in firms' average rates of return. Regression (3) combines both effects, accounting for about 15 percent of the observed variation in rates of return. The correlation coefficients

[9] The term *profit* as used here is roughly equivalent to net business income, i.e., the difference between accounting revenues and costs. To adjust for differences in firm size, profit is usually expressed as a percentage of some base. The choice of a profit base is important for some industries. Aerospace profits, for example, when measured as a percentage of assets rather than net worth (Refs. 1, 29) differ substantially in rank compared with other groups. Among the many possible measures (Refs. 2, 15, 33), rate of return on net worth appears the most appropriate for studies of the risk-profit relationship.

[10] The mean may not be an appropriate proxy for managers' anticipations if earnings are serially correlated. In such a case, earnings can be predicted from knowledge of the autoregressive structure so that computing moments about the mean would tend to overstate the firm's risk exposure. To compensate for this possibility, we adjusted each firm's earnings to remove any trend effect and then tested for autocorrelation using the Durbin-Watson statistic (Ref. 5). Evidence of positive serial correlation was found for nine of the firms, and they were removed from the sample.

Following the convention established in Refs. 4 and 30, we used standard deviation, rather than variance, as a measure of dispersion. Also, since we are concerned with the ability of firms to predict profit rates, the rates of return are unweighted.

are low, but the estimates of b_1 and b_2 are statistically significant at the .01

and .10 levels, respectively. Moreover, the signs of these coefficients agree with the theoretical model. Thus, these results lend support to the hypothesis that rates of return should be larger for firms with greater risk exposure.

TABLE 1.—*Relationship between risk and rate of return*

Regression	Intercept, r_0	Standard deviation, b_1	Skewness, b_2	R^2	F
(1)	0.0923	1.0452 (.3319)	----------	0.1141	9.914 (1,77)
(2)	.1488	----------	−0.0159 (.0095)	.0350	2.794 (1,77)
(3)	.0969	1.0181 (.3264)	−.0193 (.0099)	.1560	7.024 (2,76)

The value of the intercept, r_0, implies an expected rate of return of 9.7 percent for firms with no risk. This is not a "risk-free" rate of return, however, at least not in the sense that yields on government bonds sometimes are so interpreted. The intercept, r_0, is the result of extrapolating the risk-profit relationship to the axis, and so it is the repository for all influences on profits not encompassed by the standard deviation and skewness coefficients. These implicit influences may contain elements that might be regarded as risk factors. Moreover, since no firm in the sample was without some degree of standard deviation and skewness, a risk-free rate of return cannot be directly observed. For these reasons, r_0 will be referred to as the "risk-adjusted" rate of return; it is the expected profit rate after allowing for the influence of earnings variability.

The low R^2 values indicate that, although there is some relationship between average rates of return and the measures of risk exposure, other factors account for the major part of the observed differences in rates of return. Differences in market structure, technology, managerial ability, capital structure and similar broad industry effects could produce substantial industry earnings differentials.

To account for differences in industry characteristics, dummy variables are introduced into the regression to capture the influence of industry-specific factors. We assume that the relationship between rate of return and the risk variables is not influenced by group membership; therefore, the risk coefficients remain the same for all firms. Thus, the premium for risk exposure does not reflect other industry characteristics. The relationship becomes

$$\bar{r}_{ij}=C_j+b_1\sigma_{ij}+b_2 S_{ij}. \tag{4}$$

where C_j is the intercept for firms in industry j, and all other variables are as previously defined except for the addition of a subscript designating industry membership.

Estimates for b_1, b_2, and C_j appear in Table 2. The estimates for the risk coefficients, b_1 and b_2 are significant at the .01 and .05 levels, respectively, and their signs again agree with expectations. The estimates for C_j are all significant at the .05 level. The inclusion of industry variables considerably improves the explanatory power of the model; nearly half of the variation in observed rates of return is explained by the independent variables.

TABLE 2.—*Risk-rate of return relationship with industry effects*

Standard deviation, b_1	Skewness, b_2	R^2	Industry effects	
			C_j	Industry
1.0043 (.3648)	−0.0153 (.0071)	0.4936	0.1664	Drugs.
			.1335	Aerospace.
			.1131	Chemicals.
			.1026	Petroleum.
			.1021	Rubber.
			.0915	Food.
			.0857	Electrical machines.
			.0754	Automotive.
			.0724	Office machines.
			.0703	Steel.
			.0594	Textiles.

The C_j estimates are especially interesting. C_j is the jth industry's average rate of return after allowing for the influence of risk on the earnings of each of the firms in that group. In short, C_j is the average risk-adjusted rates of return. Interpreting it in this fashion permits computation of a set of average-risk premiums. This computation is the difference between the observed average rate of return for each group and its risk-adjusted rate. See Table 3 for the estimates.

Average risk premiums vary substantially, suggesting important differences in risk exposure among industries. The risk premium accounts for a sizeable part of the observed rate of return in some groups. For example, the average risk premium for the automotive and office machine groups is 7.2 and 6.8 percent, respectively. In contrast, the average risk premium is only 1.2 percent for steel firms and .75 percent in the rubber group, indicating that average risk exposure for firms in these two industries is nominal.

Table 3 also illustrates that adjustment of average industry earnings to reflect differences in firms' risk exposure narrows interindustry earnings differentials. Nonetheless, significant differences in average risk-adjusted rates of return remain. The risk-adjusted rates for the drug, aerospace, and chemical groups, for example, are noticeably larger than for the remaining groups.

TABLE 3.—*Observed and adjusted industry rates of return and average industry risk premiums*

Industry group	Average observed rate of return	Rank	Risk-adjusted rate of return	Rank	Average risk premium
Drugs	0.1832	1	0.1664	1	0.0168
Aerospace	.1570	2	.1335	2	.0245
Chemicals	.1409	4	.1131	3	.0278
Petroleum	.1147	7	.1026	4	.0121
Rubber	.1096	8	.1021	5	.0075
Food	.1072	9	.0915	6	.0157
Electrical machines	.1196	6	.0857	7	.0039
Automotive	.1477	3	.0754	8	.0723
Office machines	.1408	5	.0724	9	.0684
Steel	.0825	10	.0703	10	.0122
Textiles	.0789	11	.0594	11	.0195

When interpreting these results, it is important to remember that the model yields risk-comparable rates of return rather than estimates of "adequate" or "required" profits.[11] The accounting profits reported on corporate income statements include a variety of functional returns and nonfunctional rents, and a study of profit-adequacy should adjust for all such elements. This study has abstracted one element of accounting profits—the risk premium—and has adjusted the firm rates of return to make them comparable in this dimension. Normative judgments on the basis of risk-adjusted profit rates are therefore hazardous.

IV. AN ALTERNATIVE MEASURE OF RISK

Another measure of risk—the standard deviation of firms' average rates of return on an industry-wide average—was suggested by Cootner and Holland in their investigation of risks and profits (Ref. 4). The authors' economic rationale for this approach is quoted below:

"If we assume that an entrepreneur entering an industry is purchasing a proportionate share of the experience of every firm in the industry, then it would seem that the dispersion of company rates of return around the average rate of return for the industry in which they belong is an indication of the riskiness of an investment in that industry. Since the standard deviation of such rates of return indicates to an investor the likelihood that he would fare differently from the industry average, we would expect that if executives were risk-averters, large standard deviations would require high average rates of return to attract investment." [12]

[11] See Ref. 4 for a concept of risk-comparable profits. References 15, 16, and 19, which consider the appropriateness of the aerospace rate of return, define profit "adequacy."
[12] See Ref. 4, p. 4.

Assuming that an entering firm cannot identify the factors that read to intra-industry earnings differentials, this approach measures the risk of entering an industry. It is not clear, however, why a firm already engaged in some industry should be concerned with the industry average. The firm's own history would seem a better guide to the future than the overall industry experience. Nonetheless, in order to compare the intraindustry-dispersion approach to risk with the approach used in the previous model, average risk-adjusted rates of return have been computed using the following equation to measure standard deviation:

$$\sigma_j = \left[\frac{\sum_{t=1}^{n} \sum_{i=1}^{m} (r_{it} - R_j)^2}{nm - 1} \right]^{1/2} \tag{5}$$

where

σ_j = standard deviation of firm rates of return about the industry average, industry j;
R_j = average rate of return on net worth in industry j;
r_{it} = rate of return for firm i during year t;
n = number of years in sample;
m = number of firms in industry j.

The relationship between risk and rate of return becomes

$$R_j = R_o + b\sigma_j \tag{6}$$

where R_o = intercept, and b is the marginal effect of intraindustry dispersion on average industry rates of return. Estimates of these terms for the 11-industry sample are: [13]

$$R_j = 6.979 + 1.084\ \sigma_j \quad R^2 = 0.734. \tag{7}$$
$$(0.223)$$

The average risk-adjusted rate of return for each industry, R_j^*, can be computed from

$$R_j^* = R_j - b\sigma_j. \tag{8}$$

Table 4 compares these estimates of the average risk-adjusted rates of return with those obtained above using the previous measure of risk exposure. For most industry groups, the risk-adjusted rates of return are not greatly affected by the choice of a measure of risk. Two that differ substantially, however, are the drug and aerospace groups. The intraindustry-dispersion measure results in a disk premium of about 8 percent for both groups. Measuring risk by temporal earnings variability results in risk premiums of 1.6 and 2.5 percent, respectively.

TABLE 4.—*Average risk-adjusted rates of return and risk premiums* [1]

Industry	Risk-adjusted rates of return		Risk premiums	
	Intraindustry dispersion	Firm-temporal dispersion	Intraindustry dispersion	Firm-temporal dispersion
Drugs	0.1042	0.1664	0.0790	0.0168
Aerospace	.0772	.1335	.0808	.0245
Chemicals	.0995	.1131	.0414	.0278
Office machinery	.0605	.0724	.0803	.0684
Electrical machinery	.0596	.0857	.0509	.0339
Petroleum	.0898	.1026	.0249	.0121
Rubber	.0791	.1021	.0305	.0075
Food	.0604	.0915	.0468	.0157
Steel	.0566	.0703	.0259	.0122
Textiles	.0487	.0594	.0302	.0195
Automotive	.0619	.0754	.0858	.0723

[1] The firm-temporal dispersion figures include the effects of both skewness and standard deviation of firm's earnings. The intraindustry figures reflect only the effect of standard deviation. While these alternative estimates of risk-adjusted rates of return and risk premium are not computed the same way, the figures are consistent with the conceptual basis underlying each alternative measure of risk exposure. Moreover, excluding the effect of skewness from the firm-temporal dispersion figures has little effect on the magnitude of the risk-adjusted rates of return or risk premiums.

[13] Significant at the 0.01 level.

The drug firms in the sample differ substantially in their average rates of return. Consequently, intraindustry dispersion is large. Each drug firm, however, has relatively stable earnings over time so that the standard deviation measurement about each firm's own mean is small. Aerospace firms have earnings that vary substantially both about the industry average and about their own means. Correcting for trend and autocorrelation, however, results in relatively stable earnings for each firm's own mean. Thus, the intraindustry measure of dispersion overstates the risk exposure of firms in both of these industry groups and makes estimates of risk for the constituent firms dependent upon the meaningfulness of the industry groupings.

In short, the two measures of risk exposure yield disparate results for the industry groups with the highest rates of return. The choice of an appropriate measure of risk exposure is crucial for studies of risk-comparable profits.

As a risk measure, intraindustry dispersion presents several problems. For intance, if rates of return for firms in an industry such as drugs differ substantially, an intraindustry dispersion measure will indicate a substantial degree of risk even if each firm's rate of return is stable from year to year. Conceiving risk to be the difficulty in forecasting rates of return, such a method could greatly overstate the inability of existing firms to predict their future profits.

If all firms in an industry produce similar products, compete in the same markets and, in general, face exactly the same demand and supply conditions, the intraindustry dispersion measure is perfectly appropriate. With broad industry definitions, such as those used in this paper, the constituent firms are usually differentiated—while each firm shares some common elements of risk with the other firms in its group, its peculiarities create some special risk conditions. Treating each industry group as a unit, as the intraindustry dispersion measure does, obscures these firm differences. Computing the standard deviation of profits from the firm's own mean permits the industry-risk effects to be treated separately as a residual after accounting for the elements of risk peculiar to each firm in the group.

In sum, the firm-temporal dispersion measure described in Sec. II appears to have a sounder theoretical base than the intraindustry approach to risk exposure. Nonetheless, it should be noted that the concept of risk differs somewhat between the two measures. The latter is directed toward measuring the risk of entry into an industry, while the former treats risk more generally in terms of the uncertainty of forecasting future rates of return.

V. CONCLUSIONS

Perhaps the most important conclusion is implicit. With some reasonable assumptions, significant and instructive measurements of the relationship between risk and the rate of return can be obtained. The model described in this paper permits characteristics of earnings distributions to be used in evaluating risk exposure and its influence on profits. Application of the model to a sample of firms indicates that mean rates of return are importantly affected by risk exposure as defined here. Firms with large standard deviations have higher mean profit rates, while firms with positively skewed distributions have lower profit rates. The latter are apparently risk-averse and like the chance of "long-shots."

Another conclusion, relating to the method of computing measures of risk exposure, emerges. The firm-temporal-dispersion measure appears to have a sounder theoretical rationale than the alternative intraindustry-dispersion measure. A choice between these measures attains considerable importance, as they yield widely different results for the industry groups with the two highest average profit rates—drugs and aerospace. Selection is influenced by the underlying concept of risk; the intraindustry dispersion approach relates to the risk of entry, while a more general concept of risk seems more appropriate for analysis of the influence of risk on corporate rates of return. For many industry groups, adjusting nominal profit rates for risk exposure results in considerably lower risk-adjusted profit rates. This is not true, however, of the drug and aerospace groups. Their risk premiums are very low, and they also have the highest risk-adjusted rates of return. The explanation for such profit patterns, therefore, must be sought in factors other than risk.

REFERENCES

1. *Aerospace Facts and Figures*, Aerospace Industries Association, Washington, D.C., published annually.
2. Anthony, R. N., *Management Accounting*, Richard D. Irwin, Inc., Homewood, Illinois, 1960.
3. Arditti, F. C., "Risk and the Required Rate of Return on Equity," *Journal of Finance*, Vol. 22, March 1967, pp. 19–36.
4. Cootner, P. J., and D. Holland, "Risk and Rate of Return," Massachusetts Institute of Technology, DSR Project 9565, March 1963 (Mimeographed).
5. Durbin, J., and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression" Parts I and II, *Biometrika*, 1950 and 1951.
6. Fellner, W., *Probability and Profit*, Richard D. Irwin, Inc., Homewood, Illinois, 1965.
7. Fisher, R. A., *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh, 1932.
8. Friedman, M., and L. J. Savage, "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, Vol. 56, August 1948, pp. 279–304.
9. Georgescu-Roegen, N., "Choice, Expectations and Measurability," *Quarterly Journal of Economics*, Vol. 68, November 1954, pp. 503, 534.
10. Gordon, R. A., "Enterprise, Profits and the Modern Corporation," *Explorations in Economics*, 1936, pp. 306–316; reprinted in W. Fellner and B. F. Haley (eds.), *Readings in the Theory of Income Distribution*, The Blakiston Company, Philadelphia, 1951, pp. 558–570.
11. Hart, A. G., "Risk, Uncertainty, and the Unprofitability of Compounding Probabilities," O. Lange, et al. (eds.), *Studies in Mathematical Economics and Econometrics*, The University of Chicago Press, Chicago, Illinois, 1942, pp. 110–118; reprinted in W. Fellner and B. F. Haley, *Readings in the Theory of Income Distribution*, The Blakiston Company, Philadelphia, 1951, pp. 547–557.
12. Hicks, J. R., *Value and Capital*, 2d ed., Oxford University Press, London, 1950.
13. Hirshleifer, J., "Investment Decision Under Uncertainty: Applications of the State-Preference Approach," *Quarterly Journal of Economics*, Vol. 80, May 1966, pp. 252–277.
14. Hirshleifer, J., "Investment Decision Under Uncertainty: Choice-Theoretic Approaches," *Quarterly Journal of Economics*, Vol. 79, November 1965, pp. 509–536.
15. Jacoby, N. H., and J. F. Weston, "Profit Standards for Renegotiation," in J. F. Weston (ed.), *Procurement and Profit Renegotiation*, Wadsworth Publishing Company, Inc., San Francisco, 1960, pp. 121–159.
16. Jones, D. N., "Extension of the Social Control of Utilities," *Land Economics*, Vol. 41, November 1965, pp. 297–302.
17. Knight, F. H., "Profit," *Encyclopaedia of the Social Sciences*, The Macmillan Company, New York, pp. 480–486, reprinted in W. Fellner and B. F. Haley, *Readings in the Theory of Income Distribution*, The Blakiston Company, Philadelphia, 1951, pp. 533–546.
18. Knight, F. H., *Risk, Uncertainty, and Profit*, Houghton Mifflin Company, Boston, 1921, reissued by London School of Economics, 1933.
19. Little, Arthur D., Inc., *How Sick is the Defense Industry?*, Cambridge, Massachusetts, 1963.
20. Lutz, F., and V. Lutz, *The Theory of Investment of the Firm*, Princeton University press, Princeton, New Jersey, 1951.
21. Markowitz, H. M., *Portfolio Selection*, John Wiley and Sons, New York, 1959.
22. Miller, M. H., and F. Modigliani, "Cost of Capital to the Electric Utility Industry," *American Economic Review*, Vol. 56, June 1966, pp. 334–391.
23. Mills, E. S., "Uncertainty and Price Theory," *Quarterly Journal of Economics*, Vol. 73, February 1959, pp. 116–130.
24. Modigliani, F., and M. H. Miller, "The Cost of Capital, Corporation Finance, and the Theory of Investment," *American Economic Review*, Vol. 48, June 1958, pp. 261–297.
25. Nelson, R. R., "Uncertainty, Prediction and Competitive Equilibrium," *Quarterly Journal of Economics*, Vol. 75, February 1961, pp. 41–62.

26. Pratt, J. W., "Risk Aversion in the Small and in the Large," *Econometrica*, Vol. 32, April 1964, pp. 122–136.
27. Schumpeter, J. A., *The Theory of Economic Development*, Harvard University Press, Cambridge, Massachusetts, 1934.
28. Sharpe, W. F., "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance*, Vol. 19, September 1964, pp. 425–442.
29. Stanford Research Institute, *The Industry-Government Aerospace Relationship*, Stanford Research Institute, Menlo Park, California, 1963.
30. Stigler, G. J., *Capital and Rates of Return in Manufacturing Industries*, Princeton University Press, Princeton, New Jersey, 1963.
31. Tintner, G., "A Contribution to the Non-Static Theory of Choice," *Quarterly Journal of Economics*, Vol. 51, February 1942, pp. 274–306.
32. Tobin, J., "Liquidity Preference as Behavior Towards Risk," *The Review of Economic Studies*, Vol. 25, February 1958, pp. 65–86.
33. Weston, J. F., "A Generalized Uncertainty Theory of Profit," *American Economic Review*, Vol. 40, March 1950, pp. 40–60.
34. Weston, J. F., and N. H. Jacoby, "Profit Standards," *Quarterly Journal of Economics*, Vol. 66, May 1952, pp. 224–250.

APPENDIX III—ADDITIONAL SUBMISSIONS BY THE PHARMACEUTICAL MANUFACTURERS ASSOCIATION

MAY 1, 1968.

MEMORANDUM

To: Senator Gaylord Nelson, Chairman, Monopoly Subcommittee
From: Benjamin Gordon, Staff Economist

During the course of the Subcommittee's hearings on December 19, 1967, the PMA was asked and agreed to supply certain information which was not available at the moment.

The attached material which was received by the Subcommittee on April 17, 1968, after the transcripts had already been sent to the Government Printing Office, will be included in an appendix of the printed record of the Subcommittee's hearings.

This new material was not subjected to scrutiny during a public hearing, nor was it checked for accuracy by the committee staff. A cursory examination, however, revealed that, on page 13, for example, the case of MER–29 was used to show the financial risk to the industry from unforeseen side effects of a drug. It is stated that this drug was withdrawn because of unexpected side effects.

The fact is that the company was aware of the side effects before applying for the New Drug Application and withheld this vital information from the FDA and the medical profession, as a result of which, on July 2, 1963, a special Grand Jury charged, in a twelve-count indictment, that the manufacturer of this drug had violated the Food, Drug and Cosmetic Act. On June 4, 1964, the firm was found guilty in criminal case #1211–63 in the U.S. District Court for the District of Columbia.

The PMA is well aware of the facts in this case, and the example casts doubt on the usefulness of the rest of its submission.

PHARMACEUTICAL MANUFACTURERS ASSOCIATION,
Washington, D.C., April 15, 1968.

Hon. GAYLORD NELSON,
Chairman, Monopoly Subcommittee of the Senate, Small Business Committee, U.S. Senate, Washington, D.C.

DEAR SENATOR NELSON: In the course of the testimony of witnesses on behalf of the Pharmaceutical Manufacturers Association before your Subcommittee on December 19, 1967, you or Mr. Gordon asked for certain additional material. Enclosed is a memorandum supplying such material, to the extent that we have been able to obtain it, for inclusion in the record.

Sincerely yours,

C. JOSEPH STETLER.

SUPPLEMENTAL DATA REQUESTED BY SENATOR NELSON AND MR. GORDON (PAGE REFERENCES ARE TO TEXT OF TYPED TRANSCRIPT OF DECEMBER 19, 1967 HEARING)

1. DIFFERENCES BETWEEN UNITED STATES AND CANADIAN THORAZINE PRICES (PP. 2705–09)

Senator Nelson requested an explanation of the differences, indicated in a table prepared by the subcommittee staff, in prices of chlorpromazine in the U.S. and Canada (pp. 2705–09).

Under the name of its Canadian affiliate, Poulenc Ltd., Rhone Poulenc manufactures and sells its own chlorpromazine in Canada, using the Rhone Poulenc brand name Largactil. We have been informed by Poulenc Ltd. that, after the recent removal of the Canadian manufacturer's sales tax, its price to pharmacists is $18.60 (Canadian) per 500 tablets in 25 mg. strength and $45.45 per 500 tablets in 100 mg. strength. On a per thousand tablet basis, this would correspond to $37.20 (Canadian) for 25 mg. strength and $90.90 (Canadian) in 100 mg. dosage. These prices are somewhat lower but not dramatically lower than the listed Red Book prices of $57.58 and $92.64 for Thorazine in the United States, cited in the subcommittee's table. This is not unusual: prices for some drug products are lower and in other cases are higher in Canada than prices for comparable products in the United States, depending on a number of circumstances including the place of manufacture and whether royalties are paid.

Poulenc Ltd. has advised PMA by letter that Bell-Craig, the Canadian supplier cited in the subcommittee's table, "has not been licensed by Poulenc to sell chlorpromazine." The Poulenc letter makes it clear that Bell-Craig was not using Poulenc material, adding that they "do not know the source of their material, but find it difficult to understand how Canadian material can be sold at the prices quoted."

Poulenc Ltd. has also written that unlicensed imported chlorpromazine has been sold in Canada by several companies and that a test suit is now pending against one importer alleged to be distributing such material.

The comparison of Canadian and American prices of chlorpromazine shown in the subcommittee's table does not involve like products or like situations. Thorazine and Largactil are made in different countries by different companies. Prices charged to commercial outlets in the United States and Canada by the two innovating companies do not show an extreme disparity, considering the probable differences in the level of costs and other factors.

To make a valid comparison, one would have to duplicate the situation exactly, or else find some mechanism for making comparable the markets to be compared, the company producer, and the product. This is not feasible in the case sited. But even if it had been, the significance would be limited. The Sainsbury Committee in the United Kingdom, in its 1967 report, quite correctly rejected the use of such international price comparisons as a basis for proving whether drug prices in Great Britain were reasonable. In this regard the report stated in part (p. 50, para. 168):

"There are many reasons why international comparisons of prices are extremely difficult to make and even more difficult to interpret: the use of current rates of exchange as a means of putting prices expressed in different currencies on a comparable basis gives rise to many ambiguities. There are even greater difficulties if one adopts other bases for comparisons. The great differences in habits and in per capita incomes lead to differences in elasticities of demand, and the differences in the assortment of medicines consumed in different countries reduce the significance of comparisons. For these and other reasons, we have not thought it useful to attempt an elaborate comparison of international prices, for in this field figures can easily be used to arrive at any desired result."

The hazards of such comparisons are, of course, even greater when one is considering the case of prices bid by different companies, in different countries, to different government departments; particularly since such pricing is customarily on an incremental or marginal cost basis.

1A. EXTENT OF SMITH KLINE & FRENCH RESEARCH AND DEVELOPMENT OF THORAZINE (PP. 2705–09, 2894–95)

In connection with the discussion of Thorazine prices, Mr. Cutler for PMA offered to submit information on the extent to which Smith Kline & French performed research and development work on chlorpromazine in the United States (pp. 2709, 2895).

Smith Kline & French, the United States licensee of Rhone Poulenc, invested heavily in the development and testing of chlorpromazine although the initial discovery was made by the French company. A letter dated November 21, 1961 from the Director of Rhone-Poulenc* stated:

"Upon these first experiments on a limited scale we started our approaches to prospective licensees, especially in the United States, in order to widen the field of experimentation and bring out the product on the market as soon as possible.

"Our license agreement with Smith Kline & French was entered into because this company had expressed an enthusiastic interest in clinically testing and marketing chlorpromazine. Upon the basis of what was then known of the properties, this company was ready to proceed. After the license agreement, Smith Kline & French did proceed with the clinical program. Our own clinical testing efforts on a relatively small scale had preceded those of Smith Kline & French, but otherwise clinical testing by this company was concurrent with our own clinical testing efforts. As a consequence of both their clinical testing activity and our clinical testing activity, it was confirmed that chlorpromazine had the very unusual property which we now call ataraxic effect.

"In connection with the activity of both our company and Smith Kline & French in proceeding with clinical testing programs on chlorpromazine ataraxic activity, it should be remembered that the medical profession took a most dubious view toward the ataraxic effects and their possible value. The key step in proving the merit of chlorpromazine and of bringing about its use by the medical profession was the enthusiastic exploitation of every possible chance to bring home to the medical profession the results of such tests. It is very doubtful that the huge task of experimenting chlorpromazine against such odds would have been undertaken by Smith Kline & French were it not that the further marketing of the product would be protected under patent rights to permit the financial returns of the expenses involved and of the risks."

The attached Smith Kline & French pamphlet gives further details on the cooperative research and development efforts of the French and American companies on this product.

2. SUBSTANTIAL COMPANIES NOT INCLUDED IN RISK-RETURN STUDY (2789–90)

Senator Nelson requested a list of the "substantial companies" not included in the Standard and Poor Compustat tapes on which was based the Arthur D. Little, Inc. study entitled "Risk and Return in American Industry" (p. 2789).

If the criterion for a "substantial company" in the prescription drug industry is set at an annual sales level of $30 million, then the PMA's records indicate that the Standard and Poor and, therefore, the Arthur D. Little, Inc. study omitted the following "substantial" prescription pharmaceutical manufacturing companies from the sample:

Lederle Laboratories (American Cyanamid).
E. R. Squibb & Sons (Olin Mathieson at that time).
Hoffmann-La Roche, Inc.
CIBA Pharmaceutical Company.
Geigy Pharmaceuticals.
Sandoz Pharmaceuticals.
Burroughs Wellcome & Co. (U.S.A.) Inc.
A. H. Robins Company, Inc.

There was good reason for the omission of all of these firms from the sample. The first two are divisions of larger chemical firms; no separate financial reports are available for the divisions. The next five are all subsidiaries of foreign-based firms; financial statements are not available separately for the United States operations. A. H. Robins Company became a public corporation only recently, so that data required for the Risk-Return study would not have been available for the entire time period covered.

On the other hand, the Standard and Poor tapes include one firm, Gillette, which is not a pharmaceutical manufacturer. Arthur D. Little, Inc., for purposes of its study, left Gillette in the sample to avoid any possible distortion of the data from the original Standard and Poor tape. The inclusion of Gillette, however, gives an upward bias to the average of drug industry profits, since Gillette

*_Drug Industry Antitrust Act_, Hearings before Subcommittee on Antitrust and Monopoly, Senate Judiciary Committee, on S. 1552, 1962, part 4, pp. 2157–58.

has earned a return on common stockholders' equity over twice as high as the average for the drug companies and higher than even the most profitable of the drug companies in the sample.

The twenty-nine companies used in the sample include two firms which are not PMA members. One, McKesson & Robbins, is an associate. The other, Plough, is a manufacturer of proprietary products.

3. PERCENTAGE OF DRUGS PRODUCED BY COMPANIES IN RISK-RETURN STUDY (P. 2790)

Senator Nelson asked that information be supplied as to the percentage of drug products represented by the companies included in the Standard and Poor Compustat tapes on which was based the Arthur D. Little, Inc. study entitled "Risk and Return in American Industry" (p. 2790).

Of the 29 companies included in the Standard and Poor Compustat tapes, there are 26 which are members of PMA. These 26 firms account for about 70 percent of the ethical sales of United States firms, both here and abroad, according to PMA records.

4. SPATIAL VERSUS TEMPORAL VARIANCE AS A MEASURE OF UNCERTAINTY

In discussion of the Arthur D. Little, Inc. study entitled "Risk and Return in American Industry," Mr. Plotkin indicated that Arthur D. Little, Inc. would supply a memorandum explaining "why in economic terms the temporal variance . . . is not a good measure of uncertainty . . ." (p. 2794).

It is understood that Arthur D. Little, Inc. is separately submitting, for inclusion in the record, further material on aspects of its study which have been questioned during the hearings.

5. DETAILED FINANCIAL DATA PER PRODUCT (P. 2804–05)

Mr. Gordon requested, for each drug product cited by Professor Markham as made obsolete by a new product, that data be given on the following: the investment in the product, the amount recouped, the loss if any, the amount spent on research and development, and the effect obsolescence of the product had on profits of its manufacturer (p. 2805).

We have confirmed that for the reasons given by Professor Markham (pp. 2804–05), data on research and development expenditures on individual products, whether successful or unsuccessful, are not available.

However, research and development expenditures may be subdivided according to general therapeutic categories, as illustrated by the chart on the next page, reproduced from PMA's latest annual research and development survey.

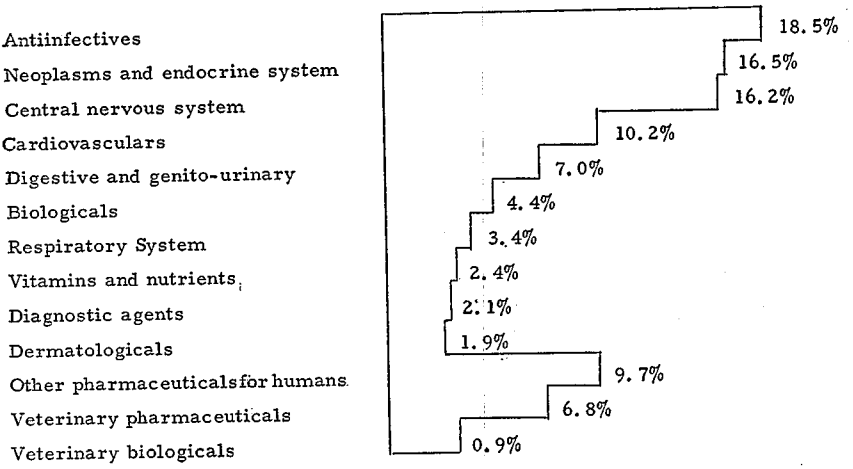
To the extent that financial information of the kind requested is available as to particular products cited, such information is given in subsequent parts of this memorandum.

APPLIED RESEARCH AND DEVELOPMENT

In 1966, the most important applied research and development field was parasitic and infective diseases. In 1965, research for the discovery and development of drugs for central nervous system and sense organs had led all others in importance with 19.0 percent of all applied research and development funds. The 1966 ratio dropped to 16.2 percent.

Cardiovascular system drugs also showed a substantial gain. Also showing relative increases in 1966 over 1965 were research expenditures for dermatologicals and drugs acting on respiratory system. However, as in 1965 more than half of the applied research and development funds were directed toward the creation and development of drugs to be effective in three classes: (1) drugs affecting parasitic and infective diseases, (2) drugs for neoplasms, endocrine system and metabolic diseases, and (3) drugs for central nervous system and sense organs.

APPLIED RESEARCH AND DEVELOPMENT EXPENDITURES BY THERAPEUTIC PURPOSE, 1966



From PMA Report: Pharmaceutical Industry, Research and Development Activity, 1966-1967 page 8.

6. COMPANY WHICH LOST TOP PLACE IN DIURETICS TO MERCK BEGINNING IN 1957 (PP. 2811–12)

Referring to Table 11A entitled "Diuretics" in the Arthur D. Little, Inc. report "Trends in Market Share for Ethical Pharmaceutical Products," Senator Nelson requested identification of the companies whose products were indicated as No. 1 and No. 5 (p. 2811). Merck & Company was tentatively identified with product No. 5. With respect to the company identified with product No. 1 (p. 2816), Senator Nelson also requested some information on the effect its drop in rank had on its profits.

It has been confirmed that product No. 5 was Merck's and that product No. 1 was Lederle's Diamox.

The precise effect of Diamox's decline upon Lederle's profits cannot be determined since figures available from American Cyanamid, of which Lederle is a division, can not be broken down by product. However, data on the rise and fall of Diamox's market share are shown on pages 42 and 43 of the Arthur D. Little report.

7. EFFECT OF POLIO VACCINE INCIDENT ON CUTTER'S PROFITS (PP. 2831–32)

Senator Nelson requested information demonstrating that the decline in the profits of Cutter Laboratories beginning in 1955 was attributable to a problem that occurred with its polio vaccine (p. 2832).

It has been confirmed from Cutter Laboratories that the losses shown for Cutter in 1955 (as recorded in the Arthur D. Little study on Risk and Return) were a reflection of an incident involving defective polio vaccine, and that the record is duly shown in the annual report of the company for 1955, which showed a loss of $1,373,535 in inventory writedown and extraordinary expenditures. The annual report for 1961 sums up the further losses from the settlement of claims against the company: these payments, together with legal fees, exceeded product liability insurance by $1,564,000.

8. EXAMPLES OF RISK FACTORS (PP. 2799, 2821–22, 2841–44)

Professor Markham referred to five risk factors by which the drug industry was characterized (pp. 2799, 2821, 2822 and 2823). He gave one example of each factor. Senator Nelson requested more examples (pp. 2821–22, 2841–44).

Additional examples are submitted below to illustrate the five categories of risk listed by Professor Markham. But it should be mentioned, by way of introduction, that Dr. Markham's list of risk categories was not exhaustive. The recent disturbances which have occurred in major American cities recalls another category of risk which particularly affects manufacturers who have maintained a tradition of replacing without cost the uninsured stocks of retailers lost by fire or other disaster. We have seen that drug stores are a favorite target of the arsonists.

Some of the broad risk situations described by Professor Markham are not unique to the pharmaceutical industry. The major risk category, involving the displacement of products on grounds of quality, applies to many other industries as well. However, pharmaceutical manufacturing *is* distinctive, in our opinion, in that the product life cycle is frequently shorter than in most other industries as a result of successful research competition. Dr. W. G. Malcolm, of the American Cyanamid Co., gave a good example of this phenomenon in his description of the high obsolescence rate of antibiotics during his testimony before the Kefauver Subcommittee (Part 24, page 13637, Hearings on Administered Prices, Sept. 7, 1960).

There are, however, other risk situations which are probably distinctively characteristic of the drug industry, such as the risk of the discovery of unforeseen side effects, the risk of unforeseen abuse of the product, the risk inherent in the much greater importance (compared to other industries) of quality control, and the greater risk of significant changes in Government regulations affecting the earnings of a given manufacturer. It should be remembered also that when, for any one of these reasons, a basic pharmaceutical ingredient has to be recalled or restricted, the impact is normally felt by more than one product in a company's line; it usually involves the recall or restriction of numerous products containing the ingredient.

It should be observed that the Arthur D. Little study entitled "Trends in Market Share for Ethical Pharmaceutical Products" shows numerous shifts in product leadership for which precise reasons cannot be assigned. As pointed out above, moreover, it is seldom possible, even where the incidence of a particular risk can be identified in the case of a particular company, to identify and quantify the effect thereof upon the company's profits. However, the following examples tend to confirm Professor Markham's analysis.

(*a*) *Risk Factor No. 1.*—Development of a competing product superior to one of a company's major products causing virtually complete replacement of it in a short time (p. 2799).

The risk of product obsolescence, which is so pronounced in the drug industry, is not confined to situations involving virtually complete displacement. Significant losses may occur through the loss of a significant market share for a given product in the area of a company's concentration. As used here, the term displacement refers to the loss of a significant market share. It does not imply that the new leader monopolizes the market, even temporarily. A study of the tables in the Arthur D. Little Inc. report on *Trends in Market Share* reveals that the displaced product normally continues in the market, but with a reduced share.

As Professor Markham testified, the experience of Lederle and Merck in the diuretics market is a good example of this type of risk.

Lederle-Diamox.—Diamox, Lederle Laboratories' brand of acetazolamide, presents an illustration of the effects of new product development on market position in the drug industry. In addition, it is a classic example of the significance of "molecular modification" in the development of important new drugs. Diamox is a member of the sulfonamide family of chemicals and was developed after it was seen that sulfonamides used as antibacterial agents had slight diuretic properties. Diuresis could only be produced, however, through use of very high dosages of these antibacterial sulfonamides.

In 1950 Lederle scientists discovered a series of sulfonamides that proved interesting as diuretics and—these agents had relatively low toxicity. Their work culminated in the introduction of Diamox in 1953. The drug offered the advantage of oral administration, proved safer than the older diuretics based on mercury, and soon enjoyed substantial success, achieving leadership in the diuretic prescription market in 1956. At that time, Diamox prescriptions accounted for 53 percent of all diuretic prescriptions. It was primarily used in the treatment of excessive body fluid retention associated with congestive heart failure.

During the period of Diamox development and early use, other firms, most notably Merck Sharp & Dohme, had undertaken research programs of their own aimed at finding sulfonamides that offered diuretic properties. Actually, the Merck program had begun in 1943, as a basic research effort to learn more about the kidney. After 12 years of investigation the firm discovered chlorothiazide; following three additional years of trial, the product was introduced for physician prescribing in 1957 as Diuril.

Diuril's success was almost immediate—by 1958, it accounted for 72 percent of diuretic prescriptions, while Diamox dropped to only 17 percent. In the years since then, Lederle has been unsuccessful in finding a product capable of recapturing the share of the diuretic market its Diamox once held. Indeed, by 1965, the Diamox share had fallen to less than 4 percent of the diuretic market. Diamox has continued to enjoy favor for use in treatment of various conditions, primarily because of its value in controlling the intraocular pressure found in glaucoma; it also has proven useful in certain forms of epilepsy. Its earlier significant share of diuretic prescriptions was lost to Diuril and the several other thiazide products introduced by various firms since 1958.

As mentioned above, financial data on Lederle Laboratories, a division of American Cyanamid Co. are not published separately.

Merck-Cortisone.—The experience of Merck & Co. with cortisone provides another outstanding example of how costly research, resulting in a major scientific breakthrough, may be followed by displacement by a competing product; causing the loss of expected financial rewards. The story is described in full detail in the 1959 hearings before the Senate Subcommittee on Antitrust and Monopoly (Part 14, p. 8013ff). Merck's interest in steroid chemistry dates from 1933. Its synthesis of cortisone, whose usefulness in rheumatoid arthritis was first demonstrated in the fall of 1948, merits the word "breakthrough."

In 1949, Merck felt that the demands of large-scale production were too difficult for one company to try to meet the public need for so major a new product. Its Board of Directors adopted a statement of policy designed "to make cortisone and other related drugs available to the public as quickly as possible in necessary quantities and at reasonable prices." To that end, Merck made it possible for its competitors to enter the cortisone market, and many firms did so. In addition, a number of firms, including Merck, undertook major competitive research programs to find steroids better than cortisone. The competition was extensive and successful.

The results of these events clearly were beneficial to patients. Major corticosteroid price reductions took place, and major steroid innovations were introduced. As for Merck: within a few years its share of the cortisone market was virtually cut in half.

In the next few years Schering introduced prednisone and other firms, among them Squibb, Upjohn and Lederle, entered the market in the prednisone family, triamcinolone, and other cortisone-related agents. Even though Merck brought out its own versions of prednisone and prednisolone, it nevertheless wound up in 1958—a decade after it had virtually created the steroid drug market—with only 17 percent of that market. The decline was even more dramatic in terms of corticosteroid plain tablets. Merck's brand of cortisone, the original steroid, dropped from 100 percent of all new steroid prescriptions written in 1950 to 3 percent in 1956 and to less than 1 percent in 1958.

It is difficult to lay changes in profit rates for a firm as diverse as Merck on the doorstep of any single product class; nevertheless, it is clear that cortisone reverses contributed to the drop in Merck profits in the early 1950's. Whereas the company had been enjoying profits on sales in the area of 10½ percent since the war, it realized only 7.9 percent in 1952 and declined further to 7.4 percent in 1953.

It would be misleading to suggest that this decline in earnings was attributable solely to the cortisone experience. In fact, other factors were involved, including drops in antibiotic and other prices. But it seems clear that the displacement of Merck corticosteroids after such an expensive research effort was an important contributing factor in the loss of earnings.

As is well known, this particular company has one of the most enviable research records of any in the world. During the middle fifties and thereafter, it continued to expand an already broad commitment of medical research, despite the competitive "squeeze." Unquestionably, Merck's introduction later in that decade of significant new products helped the company regain and improve upon its former profit position.

(*b*) *Risk Factor No. 2.*—Discovery of unanticipated side effects leading to immediate limiting of indications (p. 2817).

Professor Markham gave as an example of this second risk factor the chloramphenicol problems of Parke, Davis & Company in 1952 (pp. 2817–2820).

SKF-Parnate.—The displacement of a commercially successful product may be due to the discovery of some side effect which, because of the rarity of its occurrence, was not uncovered at the time the drug was undergoing clinical trials. An example of the financial impact of the discovery of unanticipated side effects may be seen in the case of Parnate, Smith, Kline & French's antidepressant. In 1961 after gaining knowledge that the product could cause undesirable side effects in patients who ate cheese while taking the drug, the FDA required the removal of the product from the market. It has since been restored by the FDA on the grounds that the benefits outweigh the dangers when physicians and patients are appropriately warned. Although it was removed from the market for only a brief period of a few months, the incident cost Smith, Kline & French a loss of 80 percent in sales of the product for the year.

An example from a small pharmaceutical company* is provided in the case of Imferon (iron dextran injection) which is marketed by Lakeside Laboratories, Inc. Imferon was first made available in the U.S. in the 1950's. The 1960–1961 edition of *Drugs of Choice* (Walter Modell, M.D., editor) carries the comment (page 637): "With the advent of iron-dextran (Imferon) for intramuscular injection, parenteral administration of iron has become a safe, highly efficient procedure. The preparation contains 50 mg. of iron per milliliter of solution and, aside from occasional local tenderness and temporary staining of skin at the site of injection, appears to be free from untoward reactions." More recently in an article entitled "Iron Deficiency in the U.S." (JAMA, February 5, 1968), Imferon was reported to be the iron preparation of choice for parenteral administration.

Because Imferon possesses such a low toxicity, the product was utilized in a number of research studies designed to determine the effects of overloading in animals. During the course of these researches, one of the investigators, using repeated injections of massive doses at the same site, reported a sarcoma at the injection site. Another report also appeared concerning a possible cancer in a woman who had received Imferon three years earlier. While it was uncertain that the possible cancer was caused by Imferon, the company, on the basis of this slight evidence, was requested by the FDA to show cause why the product should not be withdrawn from the market. Lakeside Laboratories, Inc. voluntarily decided in early 1960 to recall the product from the market at a very considerable cost. After conducting additional expensive and costly research Lakeside was able to demonstrate that "the risk of carcinogenesis, if any in man . . . appears to be extremely small".** The company then was again permitted to market Imferon approximately 20 months following its withdrawal from the market.

With withdrawal of Imferon was the major cause of drastic drops in both sales and profits of the company. In 1957, the year Imferon was introduced, sales of Lakeside were approximately $6,800,000 and profits approximately $620,000, while in the calendar year 1960 sales had dropped to approximately $4,200,000 and there was a loss amounting to approximately $540,000. In fact, it was not until 1963 when Imferon sales again approached the 1958 level, that Lakeside was able to show a sales profit.

Following the thalidomide tragedy in Europe, a number of different drug products were restricted or removed from the market for fear of teratogenic side effects. Meclizine hydrochloride U.S.P., an anti-nauseant compound contained in several drug products, was restricted for this reason. Other side effects include photosensitivity, which caused the restriction on the use of dimethyl chlortetracycline. Streptomycin has been restricted and dihydro-streptomycin taken off the market because of adverse reactions affecting the nervous system.

(*c*) *Risk Factor No. 3.*—Risk that drug product may be abused and require control or withdrawal (p. 2821).

Professor Markham mentioned the risk that a drug product may be misused in ways creating significant social problems leading to limitations or marketing or possible removal from the market. The example he gave was meprobamate

*Which has been owned by Colgate-Palmolive Company since February 1960.

***Physicians' Desk Reference,* 22nd edition, Medical Economics Inc. Oradell, New Jersey, 1967, p. 715.

which the Food and Drug Administration had ruled to be subject to the restrictions under the Drug Abuse Control Act.

The most dangerous drugs have been withdrawn from the market or not permitted to be marketed in the first place. Other drugs, addictive derivatives of opium, have been placed under the strict controls of the Narcotics Bureau. But there is an important third category of drugs, including the barbiturates and amphetamines, which are of proven value and far less dangerous, which have been found only recently to be subject to abuse and have hence been made subject to the Drug Abuse Control Act.

The removal of a product from non-prescription status, or if already a legend drug the placing of that product under the restrictions of the Drug Abuse Control Act, is likely to have a depressing effect on sales and as such must be considered as one of the risks facing the industry. It is difficult and often impossible to quantify the loss in such cases. Smith, Kline & French has estimated the loss of sales on its amphetamine products due to the new Drug Abuse Control Act restrictions at about 20 percent for this important category of its product line.

(*d*) *Risk Factor No. 4.*—Risk of withdrawal or restriction of products pending additional evidence of safety and efficacy (p. 2821).

It is difficult in many cases to determine whether a given example falls under this heading or under the category of risk from unforeseen side effects. The example cited by Professor Markham—MER 29—was a case in which a drug was withdrawn on the basis of unexpected side effects. The example of Imferon, mentioned above, is another case in point.

It is too early to determine the full effects of the current efficacy review of pre-1962 drug products. The most important category of products so far affected is that of the bioflavinoids. For some companies, especially smaller ones specializing in vitamins, this decision by the FDA has had a serious financial impact.

(*e*) *Risk Factor No. 5.*—Risk arising from a problem of quality control (pp. 2823–24).

Quality pharmaceutical houses maintain extensive and elaborate control laboratories and procedures. Such firms strive for perfection constantly, within reasonable economic cost limits, but absolute "zero-defects" has never proved possible in any system yet, despite highly sophisticated control mechanisms coupled with human judgment. With billions of tablets produced annually both machine and human error may occur—and when it does, it may be exceedingly costly. Because of the nature of the product, the effect of such error can be far more serious, both for the consumer and for the manufacturer, than in other types of manufacturing. An error in labeling, for instance, may cause considerable danger or it may be relatively innocuous, but the recall of products resulting from such mistakes usually is very costly.

An example of this type of risk is the Cutter vaccine incident cited by Professor Markham. See Point 7 above.

Under this same heading of risk arising from quality control problems, mention should be made of the danger of the contamination of one product by remote traces of another. This situation, leading to the recall of products from the market, can arise, of course, from the fallible human element, despite all the safeguards established under the high standards which the pharmaceutical companies set for themselves. But it can also occur as a result of changes, sometimes unexpected, in Government standards or regulations, so that products which formerly fell well within the prescribed tolerance limits are suddenly subject to recall, with little or no advanced notice and with all the loss, both financially and in prestige, which such recalls entail.

Industry witnesses have already testified about the unforeseen investments required as a result of more stringent regulations being imposed to eliminate even a trace of penicillin contamination of other products. In at least one case this compelled a major company to build an entirely new factory some miles distant from its other plants. Many have had to incur substantial costs of pulling products back from wholesalers and retailers. The quality control problem may even be due to error on the part of the authorities. The February 12, 1968 issue of the F–D–C Reports (Pink Sheet) cited the example of an FDA recall of 27 million tablets of a heart drug which appears to have been due to a technical error in the assaying techniques used by the FDA's inspection authorities. The USP, whose assay directions were used in this case, announced that discovery of an "un-

recognized discrepancy" in these directions "led to the conclusion that the tablets probably were not actually below the required potency." There is really nothing that a manufacturer can do to protect himself from this sort of risk.

9. LIST OF PATENTED PRODUCTS USED IN PRICE INDEX (PAGE 2865)

Dr. Firestone indicated his willingness to supply the list of products used in the preparation of the special price index for patented products. To avoid possible misunderstandings of the composition and significance of this material, we believe it would be preferable if Mr. Firestone were invited to come to Washington to discuss his worksheets and methods with the subcommittee staff.

O