ESTIMATING THE EFFECTS OF REGULATION ON INNOVATION: AN INTERNATIONAL COMPARATIVE ANALYSIS OF THE PHARMACEUTICAL INDUSTRY*

HENRY G. GRABOWSKI, JOHN M. VERNON,
and
LACY GLENN THOMAS
Duke University

INNOVATION in the U.S. ethical drug industry in recent years has been characterized by a number of adverse developments. In particular, there has been a sharp decline in the rate of new product introductions and the incentive for engaging in research and development (R & D) activity has been negatively influenced by rapid increases in the costs and risks of developing new products. While there is little debate about the existence of these adverse trends, there is considerable controversy about the factors producing them.

Briefly, we list below five hypotheses that have been discussed as explanations for the declining rate of innovation.

(1) Tighter regulation of the industry by the Food and Drug Administration (FDA) has been largely responsible for the declining rate of innovation.
(2) The decline is illusory—while there has been a decline in the *total* number of new drugs being introduced, the number of "important" new drugs introduced annually has not declined.
(3) There has been a "depletion of research opportunities" brought about by the rapid rate of new drug development in the 1950s.
(4) The tragic thalidomide episode in the early 1960s made drug firms and physicians much more cautious in their decisions concerning the marketing and prescribing of new drugs.
(5) Advances in pharmacological science have led to increased safety testing and, therefore, higher costs of developing new drugs.

In this paper, we present some new evidence on these hypotheses. Our

* We are grateful for the comments we received on a preliminary version of this paper presented at the Third American University Seminar on Pharmaceutical Public Policy Issues. In addition, we received helpful comments from Sam Peltzman, Dudley Wallace, and Oliver Williamson. The research was supported by the National Science Foundation, Division of Policy Research and Analysis.

new evidence is based primarily on a comparative analysis of developments in the United States and United Kingdom. In particular, we attempt to separate the impact of increased regulatory controls in the United States (stemming from the 1962 amendments to the 1938 Federal Food, Drug, and Cosmetic Act)[1] from other factors by using the U.K. industry as a control. Since firms in the latter country have been governed by a very different regulatory system but are similar to U.S. firms in most other ways, we feel that comparative analysis is a very fruitful way of approaching this question.

The paper has the following plan. First, as background to our analysis, we briefly describe the structural changes that have characterized new product innovation in ethical drugs, as well as the hypothesized relations which account for these trends. We then review two past empirical studies that have attempted to explain the most important and controversial of such structural changes: declining levels of new product introductions in the United States. Finally, a model previously developed by Martin Baily[2] is reformulated and employed in a comparative analysis of the U.S. and U.K. industries.

I. Structural Changes in Pharmaceutical Innovation: Trends and Hypotheses

Evidence from a number of studies indicates that the American pharmaceutical industry has undergone some fundamental shifts in innovational structure and performance over recent years. This section briefly documents these basic trends and more systematically considers the proliferating hypotheses which have been advanced to explain these structural changes.

A. *Trends in Pharmaceutical Innovation*

In the post-1962 period, the U.S. pharmaceutical industry has experienced the following.

i) *Declining Rates of New Product Introductions.* This decline is illustrated in Figure I. It shows the total new chemical entities (NCEs) introduced annually into the United States over the period 1954-1974, as well as the subset of each year's introductions that were discovered in the United States by the pharmaceutical industry.[3] NCEs are the most important cate-

[1] Federal Food, Drug, and Cosmetic Act of 1938, 52 Stat. 1040, c. 675 as amended by Pub. L. No. 80-625, 21 U.S.C. §§ 1-517 (1964).

[2] Martin N. Baily, Research and Development Costs and Returns: The U.S. Pharmaceutical Industry, 80 J. Pol. Econ. 70 (1972).

[3] Data on NCEs and their years of introduction were obtained from Paul de Haen, Inc. See note 54 *infra*. Biologicals and diagnostics were deleted from the analysis. Information on the country of discovery was also obtained from de Haen, as well as supplementary sources. An NCE is regarded as discovered in a particular country if the research laboratory producing the

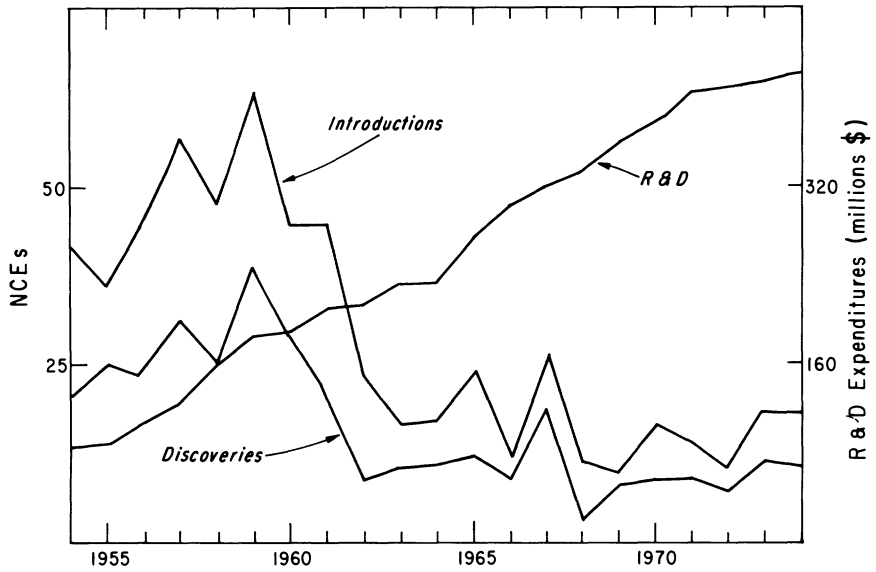

FIGURE I

Introductions and Discoveries of New Chemical Entities by Domestic Firms and Constant (1958) Dollar Expenditures on Pharmaceutical Research and Development, the United States (1954-1974).

gory of new products because they represent compounds not previously marketed and include all significant new therapeutic advances. Thus NCEs form a reasonable index of innovative output. Other new products involve combinations of existing products, new dosage forms, or new brand names.

In Table 1 data on NCE introductions are grouped into five-year periods beginning in 1957.[4] The table shows that the rate of introductions over the most recent five-year period is less than one-third the rate prevailing in a similar period a decade ago. The third column of Table 1, which shows the total market shares captured by new NCEs over these three periods, underscores the extent to which new product innovation has declined as a competitive factor in the ethical drug market.

ii) *Increasing Costs of Innovation.* Over the same time frame in which introductions and discoveries of NCEs have significantly declined, industry R & D expenditures have increased severalfold. These trends imply a rather

---

entity was located in that country, irrespective of the nationality of laboratory ownership. See the Appendix for details on the procedures used in the text in this regard.

[4] The choice of period here was dictated by the availability of sales data (no data were available prior to 1957) and the three-year average sales measure employed in Table 1. The sales data were obtained from Intercontinental Medical Statistics, Inc. See note 63 *infra*. The nature of these data is discussed in the Appendix.

TABLE 1
NUMBER AND SALES OF NEW CHEMICAL ENTITIES
IN THE PRE- AND POSTAMENDMENT PERIOD IN THE UNITED STATES

Period	Total Number of New Chemical Entities (NCEs)	Average Annual Sales per NCE (during first 3 years)	Sales of NCEs as a Percentage of Total Ethical Drug Sales[a]
1957-1961	233	$1,745,000.	20.0
1962-1966	93	$2,657,000.	8.6
1967-1971	76	$3,187,000.	5.5

[a] Average annual sales of all NCEs introduced during this period as a percentage of total ethical drug sales in the last year of the period.

Sources: Lists of new chemical entities in each year were obtained from Paul de Haen, Annual New Product Parade, various issues; all information on ethical drug sales were obtained from Intercontinental Medical Statistics, various years.

formidable increase in the costs of producing an NCE, an increase which has been documented in studies by Clymer, Mund, and Sarett.[5] In particular, Sarett suggests that over the decade 1962 to 1972, development costs per NCE rose from 1.2 to 11.5 million dollars.

iii) *Increasing Risks for Innovation.* In addition, there appears to be a corresponding increase in the risks and uncertainty associated with innovational activity. One measure of risk in this industry is the attrition rates for compounds that undergo clinical testing but fail to become commercial products. Clymer[6] estimates that in the 1950s, the attrition rate of drugs undergoing clinical tests was two out of three. The best estimate of the current situation appears to be that less than one of every ten new compounds entering clinical trials become new products.[7]

In short, the decline in new product outputs in the drug industry has been accompanied by a number of adverse structural trends on the input side of the innovational process. Total development time and costs have increased severalfold. Furthermore, innovation has become subject to greater risks and uncertainty. These adverse structural trends in both innovational inputs and outputs appear related to more fundamental underlying changes in the

---

[5] Harold A. Clymer, The Changing Costs and Risks of Pharmaceutical Innovation, in The Economics of Drug Innovation 109 (Joseph D. Cooper ed. 1970); Vernon A. Mund, The Return on Investment of the Innovative Pharmaceutical Firm, in the Economics of Drug Innovation 125 (Joseph D. Cooper ed. 1970); L. H. Sarett, FDA Regulations and Their Influence on Future R & D, 17 Int'l J. Research Management 18 (1974).

[6] Harold A. Clymer, *supra* note 5, at 152.

[7] In particular, Louis Lasagna & William M. Wardell, The Rate of New Drug Discovery, in Drug Development and Marketing 155 (R. B. Helms ed. 1975) (Am. Enterprise Inst.), present data (from a questionnaire survey of 15 large firms accounting for 80% of U.S. research) that indicate only 7.1% of all new drug investigational plans (INDs) filed by these firms between 1963 and 1967 had become approved NCEs by April 1974 (the date of their study).

innovational process. A review of the hypothesized causes of these adverse trends follows.

B. *The Hypotheses*

i) *Increased FDA Regulation.* Of the five hypotheses mentioned in the introduction, the role of increased regulation associated with the 1962 Kefauver-Harris amendments has received the most prominent attention in explaining declining pharmaceutical innovation. The antecedent 1938 Food, Drug, and Cosmetic Act required all new drugs to undergo a premarket approval process based on safety. Under this law, the FDA also had to reject a new drug compound within a period of sixty days or the new compound was automatically approved for marketing by the manufacturer.

The 1962 Kefauver-Harris amendments extended the regulatory controls of the FDA in several ways. First, it required firms to submit documented scientific evidence on a new drug's *efficacy* as well as its safety. This led to a substantial increase in the number of tests that had to be performed and submitted to the FDA. Second, the FDA was given discretionary power over the clinical research process. Thus, prior to any testing in humans, firms must now submit a new drug investigational plan (IND) that provides the results of animal tests and plans for human testing. Third, the new regulations provided for FDA approval of advertising claims. Finally, the provision of automatic approval of a new drug application (NDA) after sixty days unless the FDA took specific action was effectively repealed.

Over the post-1962 period, therefore, there has been a significant increase in both the scope and intensity of regulatory controls on ethical drugs. As a consequence, it has been postulated that the costs of discovering and developing a new drug, along with the risks and uncertainty of drug innovation, have increased; and that this, in turn, has been a major factor in the observed decline in innovational output.

ii) *Fewer Marginal and Ineffective Drugs.* The initial response of the FDA to hypothesis (i) was to argue that the observed decline in pharmaceutical innovation is in fact illusory:

> The relevant question is not and never has been how many new drugs are marketed each year, but rather how many significant, useful and unique therapeutic entities are developed. . . . The rate of development and marketing of truly important, significant, and unique therapeutic entities in this country has remained relatively stable for the past 22 years.[8]

Unfortunately, it is difficult to substantiate this FDA claim as there is no list of important new drugs upon which there is general agreement by medi-

[8] Speech by Alexander Schmidt, The FDA Today: Critics, Congress, and Consumerism (Oct. 29, 1974 before the Nat'l Press Club, Wash., D.C.).

cal experts. Most lists from academic sources, for example, show a significant downward trend in important therapeutic advances, as does at least one prior FDA ranking of important new drugs.[9] Furthermore, measures of pharmaceutical innovation based on economic criteria strongly suggest that a significant decline in real terms has occurred. The data presented in Table 1, in particular, indicate that the total market shares captured by NCEs have declined over time in comparable fashion to the total number of NCE introductions.[10]

Sam Peltzman has analyzed a related drug quality issue as to whether the large decline in NCE introductions could be explained by fewer ineffective drugs entering the marketplace after the 1962 amendments were passed. His analysis of data from three groups of experts—hospitals, panels employed by state public-assistance agencies, and the American Medical Association's Council on Drugs—does not support this view. These data suggest only a small fraction of the pre-1962 *and* post-1962 NCE introductions could be classified as ineffective.[11]

In sum, the hypothesis that the observed decline in new product introductions has largely been concentrated in marginal or ineffective drugs is not generally supported by empirical analyses. Moreover, these data analyses show no real tendency for more recently introduced drugs to have either significantly higher average market shares or efficacy rates than those introduced in earlier periods.

iii) *Depletion of Research Opportunities.* More recently, the FDA (along with some prominent members of the biomedical community) have emphasized a very different hypothesis—that the decline in pharmaceutical innovation is real, but that it is due to a depletion of research opportunities rather than increased regulation. This hypothesis has been described by former FDA Commissioner Schmidt as follows:

[9] Henry G. Grabowski, Drug Regulation and Innovation: Empirical Evidence and Policy Options (Am. Enterprise Inst. 1976).

[10] Market measures are premised on the notion that drugs which obtain the largest shares do so because they offer consumers the most overall utility per dollar. One can argue, however, that some drugs which have important therapeutic properties, but for relatively rare diseases, will tend to obtain low market shares. In addition, market shares are presumably influenced not only by the therapeutic advance of a new drug but also by the innovating firm's market power, promotional strategies, and so forth. However, for the broad aggregate comparison presented above, these qualifications are not as important as they might be in other situations. This is because there is no reason to believe that these factors have changed markedly over time, especially not in a direction so as to produce the lower market shares for new drugs shown above. For example, it seems unlikely that the lower market shares can be plausibly accounted for by a shift toward the production of a relatively greater number of drugs for rare diseases.

[11] In particular, these data suggest the incidence of ineffective new drugs was less than 10% in the pre- and post-1962 period. Peltzman also analyzes the growth rate patterns of NCEs in the pre- and post-1962 periods and argues they also support the findings of expert evaluations in this regard. See Sam Peltzman, An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments, 81 J. Pol. Econ. 1049, 1086 (1973).

Today's world includes a great number of important therapeutic agents unknown a generation ago. These include antibiotics, antihypertensive drugs, diuretics, antipsychotic drugs, tranquilizers, cancer chemotherapeutic agents, and a host of others. . . . In many of these important drug groups there are already a large number of fairly similar drugs. As the gaps in biomedical knowledge decrease, so do the opportunities for the development of new or useful related drugs. As shown by the declining number of new single entity drugs approved in the U.S., England, France, and Germany, this is an international phenomenon. This does not reflect a loss of innovative capacity, but rather reflects the normal course of a growth industry as it becomes technologically more mature.[12]

Adherents of the research-depletion hypothesis therefore are suggesting that in many major therapeutic areas we have reached a point where the probability that a new discovery will be an advance over existing therapies is quite low. Furthermore, they argue we are on a research plateau because the major disease areas left to conquer are the ones where we have the least adequate scientific understanding of the underlying biological processes. Hence, they suggest that considerable investments of basic research may be necessary before a new cycle of increased drug discoveries is likely to occur. They further point to the lower levels of drug introductions in other developed countries (where regulation has been less stringent than the United States) as important supportive evidence that a worldwide depletion of scientific opportunities has occurred in the pharmaceutical industry.

This hypothesis has been received with considerable skepticism in many scientific quarters. Some have challenged the hypotheses on conceptual grounds.[13] Others have pointed to the vast expenditures on basic biomedical research by the National Institutes of Health and other organizations as creating a renewed pool of basic knowledge which should offset any tendency toward a depletion of opportunities from prior drug discoveries.[14]

iv) *The Consequences of Thalidomide.* In addition to increased regulation and research depletion, Lebergott has pointed to the effects of the thalidomide tragedy on the behavior and expectations of physicians and drug firms as further confounding factors. In particular, he argues:

Do any of us believe that after that catastrophe, consumers were quite as likely as before to prefer new drugs to ones tested by experience? Were physicians henceforth quite as likely to prescribe new drugs—with the prospect of acute toxicity (and

[12] Examination of the Pharmaceutical Industry, 1973-74, Part 1: Hearings on S. 3441 and S. 966 Before the Subcomm. on Health of the Comm. on Labor and Public Welfare, 93rd Cong., 2nd Sess. 272 (1973-74) (statement of Alexander Schmidt).

[13] See, for example, statements by J. E. S. Parker and Harold Demsetz in Impact of Public Policy on Drug Innovation and Pricing (S. A. Mitchell & E. A. Link eds. 1976) (Am. Enterprise Inst.).

[14] B. M. Bloom, Socially Optimal Results from Drug Research, in Impact of Public Policy on Drug Innovation and Pricing 355 (S. A. Mitchell & E. A. Link eds. 1976) (Am. Enterprise Inst.).

malpractice suits) when the one chance of 10,000 ran against them? Which of our leading pharmaceutical firms would henceforth endanger its reputation (and its entire existing product line) on behalf of a new drug on quite the same terms as it did in the days when biochemists could do no wrong? . . . . Such massive changes in the U.S. perspective on drugs—we call them shifts in both supply and demand curves—had to cut the number of more venturesome drugs put under investigation since 1962. It would have done so if the entire FDA staff had gone fishing for the next couple of years.[15]

Thus, Lebergott argues that after thalidomide strong shifts occurred in the incentives facing physicians and manufacturers, which would operate independently to increase R & D costs and lower new drug introductions. His analysis points up the difficulties in trying to identify the effects of regulatory and nonregulatory factors that changed simultaneously as a result of the thalidomide incident.

v) *Advances in Pharmacological Science.* Finally, Dr. Pettinga of Eli Lilly and others have pointed to scientific advances in pharmacological science over the past few decades as another potentially important factor. In particular, he suggests that these advances, which have made teratology and toxicological studies much more sophisticated and costly in nature, would have been incorporated into drug firm testing procedures even in the absence of regulatory requirements to do so.[16] That is, drug firms would undertake many of these tests in their own self-interest, in order to reduce the likelihood of future losses in goodwill and potential legal liabilities.

In sum, while our primary objective in this paper is to identify the effects of increased regulation on declining levels of pharmaceutical innovation, a number of plausible alternative factors to regulation must also be considered. After briefly reviewing prior empirical work in the next section, we will turn to an international comparative approach to analyze these hypotheses.

C. *Prior Empirical Work*

i) *Sam Peltzman's Study.* Sam Peltzman's cost-benefit analysis of the 1962 amendments has received considerable attention in both economic and policy circles. We shall restrict our review here to only his analysis of the effects of the amendments on the rate and character of drug innovation.[17]

[15] Competitive Problems in the Drug Industry. Part 23: Development and Marketing of Prescription Drugs. Hearings Before the Subcomm. on Monopoly of the Select Comm. on Small Business, 93rd Cong., 1st. Sess. 9843 (1973) (statement of Stanley Lebergott).

[16] See remarks of Dr. Pettinga, in Regulation, Economics, and Pharmaceutical Innovation 288 (J. D. Cooper ed. 1975).

[17] Sam Peltzman, *supra* note 11.

Peltzman employs a "demand pull" model of new drug introductions by the pharmaceutical industry.[18] In particular, the supply of new drugs in his model responds with a lag to shifts in demand side factors (for example, the number of out-of-hospital prescriptions and expenditures on physician services). The model is estimated on pre-amendment data (1948-1962) and the estimated equation is then employed to forecast what the number of NCEs would have been in the post-1962 period in the absence of regulation. The effects of the 1962 amendments are then computed as the residual difference between the predicted and actual flow of NCEs.

Using this approach, Peltzman concludes that "all of the observed difference between pre- and post-1962 NCE flows can be attributed to the 1962 amendments."[19] However, his approach never formally includes or considers any of the supply side factors in the hypotheses cited above. All of the observed residual difference after 1962 is simply assigned to increased regulation. Since this residual difference can plausibly reflect the effects of a number of the other factors cited above (that is, research depletion, changing expectations, and scientific factors), it probably encompasses various nonregulatory phenomena as well.

ii) *Martin Baily's Study.* Martin Baily employs a production function model of drug development which does try explicitly to separate the effects of regulation from the depletion of scientific opportunities. He postulates that the number of new chemical entities introduced by the industry in any period is a function of lagged-industry R & D expenditures and that both regulation and research depletion operate to shift this R & D production function over time.

After experimenting with various functional forms and distributed lag relations, he estimates the following production function equation using time series data for the period 1954 to 1969:[20]

$$\log\left[\frac{N_t}{E_t}\right] = \underset{(15.96)}{4.708} - \underset{(6.13)}{1.337}\, D_t - \underset{(3.71)}{0.03854}\, P_t;$$

(t-statistics in parentheses) $R^2 = .95, \rho = -.3, DW = 1.98,$ (1)

where N_t = number of NCEs introduced and discovered by U.S. firms in year t

E_t = average industry deflated R & D expenditures for ethical drugs

[18] The analysis builds on the approach of Jacob Schmookler, Invention and Economic Growth (1966), who postulated that technological innovation generally followed demand rather than vice-versa.

[19] Sam Peltzman, *supra* note 11, at 1055.

[20] Martin N. Baily, *supra* note 2, at 77.

in the United States in years $t - 4$, $t - 5$, and $t - 6$ (it is assumed there is a fixed five-year lag from R & D outlays to introduction)

D_t = a zero-one dummy variable representing the effect of regulation (it equals 0 through 1961 and 1 afterward)

$P_t = \frac{1}{7} \sum_{v=7}^{13} M_{t-v}$ where M_t is total number of new drugs introduced from all sources (this seven-year moving average of past introductions is Baily's proxy variable for depletion).

In this formulation, R & D productivity (or NCEs per dollar of R & D invested) is related in a statistically and quantitatively significant manner to proxy variables for both regulation and research depletion. For example, the estimated coefficient on D_t implies that the annual expenditures required to develop a constant number of new drugs more than tripled in the post-amendment period.[21]

The Baily model therefore appears to perform well and suggests that both the regulation and research depletion hypotheses are valid. Nevertheless, it should also be noted that this specification does embody a number of strong assumptions. First, the model implies a fixed lag as well as constant returns to scale in the relation of NCE introductions to R & D expenditures. Second, the seven-year moving average formulation for the depletion variable has a somewhat arbitrary character; it also does not formally allow for additions to the stock of knowledge. Third, the zero-one dummy variable formulation for regulatory effects imposes the same shift factor on the entire post-amendment period (rather than a differential response over time). Finally, no attempt is made to consider additional factors such as those presented in hypotheses (iv) and (v) above.[22]

[21] Baily presents the estimated regulatory effect on costs only implicitly in a table showing the annual expenditure required to develop a constant number of drugs, before and after the 1962 change in regulation. This table indicates that costs increased by a factor of 2.35 beginning in 1962. However, these cost figures confound regulatory and depletion effects, and further embody the rather dubious property that the effect of depletion on costs after 1962 has only about half the magnitude of pre-1962 effects. This property follows from the assumption that the flow of drugs from non-U.S.-industry sources is lower in the post-1962 period and Baily's formulation of the depletion variable.

The direct regulatory effect, holding depletion constant, is calculated from the coefficient on the dummy variable, which, given Baily's specification, implies an increase in costs by a factor of 3.8. Martin N. Baily, *supra* note 2.

[22] Additional Baily assumptions include: (a) All R & D expenditures are allocated to discovery and development of NCEs. To the degree that the proportion of R & D expenditures devoted to NCEs fails to exhibit systematic shifts over the period of analysis, this assumption should not affect results. It should be remembered that relative or before-and-after effects are the focus of concern. (b) The gross national product deflator adequately represents price trends for R & D

Since the Baily model was published, several years of additional data have become available. In order to test the stability of his estimated regression equation, we reestimated it using more recent data. Baily used data covering 1954-1969, while we employ data for the longer period 1954-1974. Our reestimation of the Baily model yields the following equation:

$$\log\left[\frac{N_t}{E_t}\right] = \underset{}{-0.88} - \underset{(2.40)}{2.26\, D_t} - \underset{(8.63)}{0.003\, P_t} \qquad \underset{(0.23)}{} \qquad (1')$$

$$R^2 = .88 \qquad DW = 1.60.$$

Hence, the main finding of our reanalysis is that the coefficient of the depletion variable has become statistically insignificant, though it does continue to have the expected negative sign. The explanatory power of our reestimated equation also has declined substantially from that obtained by Baily (the R^2 declined from 0.95 to 0.88). Furthermore, a number of other functional specifications were analyzed and the research depletion variable performed poorly in each instance.[23]

Thus, neither the studies of Peltzman nor Baily would seem to provide completely satisfactory approaches for isolating the effects of increased regulation on pharmaceutical innovation from other confounding factors. Although Baily's production function model does provide a conceptual basis for separating regulatory factors from other supply side factors like research depletion, his proxy variable for research depletion is obviously highly unstable when extended forward in time.

In the next section, we present our own methodological approach for empirically isolating the effects of regulation from other factors. It is based on an international comparative analysis of developments in the United States and United Kingdom which we believe offers some important advantages over the time series analysis of a single country.

II. A Comparative Analysis of the U.S. and the U.K. Industries

Under ideal laboratory conditions, one would wish to observe the behavior of innovation in the United States in two states of the world: one with the

---

inputs in the pharmaceutical industry. There is some evidence to indicate that the movements of the two trends are highly correlated so that the gross national product deflator is an adequate proxy. See Nat'l Science Foundation, NSF72-310, A Price Index for the Deflation of Academic R & D Expenditures (May 1972).

[23] In particular, we examined both the multiplicative and linear functional specification and a number of formulations that relaxed various strong assumptions embodied in equation (1) (for example, fixed lag, regulatory dummy shift variable, and so forth). These generalizations are discussed in Table 3, in the context of our international analysis. However, the research depletion variable employed by Baily was never statistically significant in any of these alternative specifications.

1962 amendments in effect and one where they were not in effect. Given the impossibility of this experiment, a "second-best" experiment would be to find another country which was as similar to the United States as possible, and in which the regulatory pattern before and after 1962 was similar to that of the United States prior to 1962. The United Kingdom appears to be the best candidate for such an experiment.

In the analysis which follows, we specifically compare changes in R & D productivity in the United States and the United Kingdom. Our ultimate objective is to analyze the effects of regulation on R & D productivity in the United States, using the United Kingdom experience as a control for non-regulatory factors.

An international comparative analysis is of course subject to some inherent problems and biases as well as advantages. In what follows, we set out an analytical strategy designed to exploit the strengths of comparative analysis while minimizing or avoiding the problems.

A. *The U.K. Regulatory Environment*

As in the case of the United States, the United Kingdom experienced some basic changes in regulatory procedures governing drugs as a result of the thalidomide incident. Prior to 1963, the laws in the United Kingdom required registration of all new drug substances with the Ministry of Health. The main control on safety, however, came into play *after* a drug was marketed. Each registered new drug was referred to a Committee of the National Health Services for classification of its therapeutic properties.[24] Their evaluation of each drug was then disseminated to physicians. Some sanctions were available to the National Health Services to discourage physicians from prescribing drugs classified as being of "unproven value."

In 1963, the Committee on Safety on Drugs was established in the United Kingdom to undertake premarket safety reviews of drugs. Hence, the U.K. system after 1963 incorporated the basic requirement of premarket safety reviews that had been in effect in the United States for many years before 1962. At the same time, the United Kingdom did not institute most of the requirements associated with the 1962 amendments. Specifically, the United Kingdom did not require formal proof of efficacy until the Medicines Act was implemented in 1971;[25] before this act, the task of evaluating a drug's efficacy was essentially left to the market mechanism. In addition, the U.K. IND procedure was on a voluntary basis until 1971. Finally, the British

[24] See W. D. Reekie, The Economics of the Pharmaceutical Industry ch. 7, at 100-12 (1975), for a more detailed discussion of this and other historical developments with respect to the U. K. regulatory system.

[25] Medicines Act, 1968, c. 67.

system apparently relied more on outside committees of medical experts and emphasized postmarket surveillance compared with the United States.[26]

Aside from these differences in regulatory procedures after 1962, the two countries share a number of important similarities. Firms in the U.K. ethical drug industry are also characterized by high levels of R & D intensity and have produced a number of important drugs adopted on a worldwide basis.[27] In addition, both countries have high standards of medical training and practice.

Firms in the U.K. ethical drug industry should also be similarly affected by the nonregulatory factors cited in hypotheses (iii) to (v) above. First, the factor receiving the most attention—research depletion—certainly should not operate only in one particular country, but should be worldwide in scope. This is especially so given the rapid diffusion of knowledge concerning new drug discoveries throughout all developed countries. Secondly, the thalidomide incident as a factor making drug firms and prescribing physicians more cautious and thereby leading to higher costs of innovation would also be expected to operate abroad as well as in the United States. Indeed, since the United Kingdom was a country directly affected by thalidomide, one might expect it to play a greater role there than in the United States. Third, technical advances in the detection of adverse effects of new drugs would also be available to foreign firms who wished to use them for reasons of self-interest in the absence of any regulatory prodding.

A comparison of the United States and the United Kingdom therefore, would seem insightful because the regulatory environment of each country after 1962 was very different in character, while the other hypothesized nonregulatory factors for the decline in innovation in the United States would tend to operate in a similar (but not necessarily identical) manner across the two countries. Two basic problems do arise, however, which must be considered: first, the U.K. regulatory environment has not been static during the period of analysis, but rather has also experienced regulatory change, culminating in the important Medicines Act of 1971; second, there are multinational linkages across the two countries.

To deal with the former problem we will structure our analysis as follows. First, to avoid confounding the effects of depletion, thalidomide, and techni-

[26] Derrick Dunlop, The British System of Drug Regulation, in Regulating New Drugs 229 (Richard L. Landau ed. 1973). For a more detailed comparison of the two systems which reaches similar conclusions, see Louis Lasagna & William M. Wardell, *supra* note 7, Part II at 51. In particular, see ch. 10, at 109-23, for a further discussion and analysis of U.K. developments since enactment of the Medicines Act.

[27] See the comparative analysis of innovational outputs in G. Teeling-Smith, Comparative International Sources of Innovation, in Regulation, Economics, and Pharmaceutical Innovation 57 (J. D. Cooper ed. 1975); and also the material in W. D. Reekie, *supra* note 24, at 50-70 & 84-99.

cal change with the regulatory effects associated with the Medicines Act, we will focus on the period prior to 1971 in the United Kingdom. Secondly, we will make the strong assumption that all variations in U.K. trends in R & D productivity before 1971 are due to nonregulatory factors.[28] The other major U.K. regulatory change occurred, as discussed above, in 1963. In order to gauge the significance of this regulatory change for U.K. rates of innovation, we regressed R & D productivity of the United Kingdom on time and an intercept dummy for 1962 and 1963. These failed to yield statistically significant coefficients on the regulatory shift dummies, even at the 10 per cent level.[29] This is in sharp contrast to the U.S. situation and suggests the regulatory changes enacted in 1963 in the United Kingdom had far less impact on innovation in that country compared to the effects in the United States of the 1962 Kefauver amendments.

Nevertheless, there may be significant negative side effects of increased U.K. regulation on R & D productivity over this period that are not adequately captured in this model. To the extent that this is so, our strong assumption that all of the observed U.K. decline in R & D productivity before 1971 is due to nonregulatory factors will impart a *conservative* bias to our estimates of regulatory effects in the United States (since we employ these U.K. trends in innovation as a control for nonregulatory factors in the United States).

We will follow the general strategy in this paper of consciously structuring our analysis so that errors and biases operate to yield an *underestimate* of the effects of regulation on innovation.

[28] It is recognized that additional health policy changes occurred in the United Kingdom during the period of analysis. For example, beginning in 1961, the Ministry of Health was empowered to negotiate price directly on any patented drug with large sales, and the prices for such drugs repeatedly changed. (M. A. Shankerman, Common Costs in Pharmaceutical Research and Development: Implications for Direct Price Regulation, in Impact of Public Policy on Drug Innovation and Pricing 3 (S. A. Mitchell & E. A. Links eds. 1976). Quite probably these alterations of policy affected the incentives for U.K. pharmaceutical firms to invest in R & D activities. However, there is little reason to believe that policy changes other than those occurring in 1963 and 1971 and discussed above would affect the productivity of whatever R & D expenditures were undertaken. And it is only productivity which will be an object of analysis here.

[29] The least squares regression equations for the U.K., 1960 to 1970, using the intercept dummy in 1963 (D_t) were:

$$\operatorname{Log}\left(\frac{N_t}{E_t}\right) = \underset{}{1.19} \; \underset{(3.19)}{-} \; \underset{(1.14)}{.35 D_t} - \underset{(2.62)}{.11 T_t}$$

$$R^2 = .72 \qquad \rho = -.55 \qquad F = 9.57 \qquad DW = 2.48$$

$$\operatorname{Log}\left(\frac{N_t}{E_t}\right) = \underset{}{3.24} \; \underset{(2.59)}{-} \; \underset{(.69)}{.25 D_t} - \underset{(2.37)}{1.41 \log T_t}$$

$$R^2 = .71 \qquad \rho = .53 \qquad F = 8.49 \qquad DW = 2.43.$$

A second class of problems which arise in an international comparative analysis are associated with multinational linkages between the U.K. and the U.S. industries. An outline of these problems and a comparable strategy for dealing with them is presented in the section which follows.

B. *The Problems Posed by Multinational Interdependence*

In Figure II, we present trends on total NCE introductions in the United Kingdom, the subset of NCE introductions discovered by the U.K. pharmaceutical industry, and this industry's R & D expenditures on ethical drugs for the period 1960-1974.[30] Clearly the trends depicted for the United Kingdom in Figure II are qualitatively similar in nature to those shown for the United States in Figure I. That is, total NCE introductions and discoveries in each country decline over time, while R & D expenditures increase.

FDA Commissioner Schmidt has argued that the downward trend on total NCE introductions in the United Kingdom (and other Western European countries)—paralleling the U.S. trend—provides evidence for a worldwide

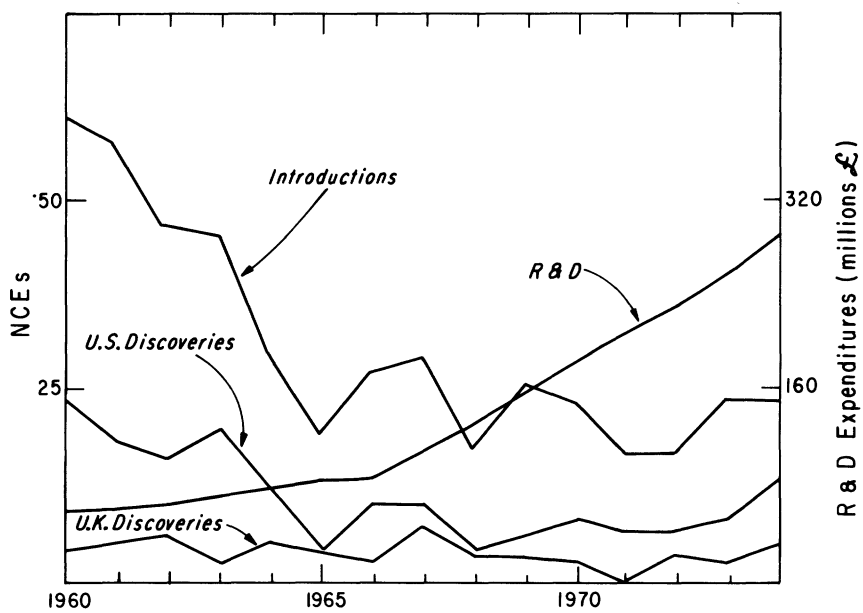

FIGURE II

Introductions of New Chemical Entities (Total Discoveries by U.K. Firms and by U.S. Firms) and Constant (1958) Pound Expenditures on Pharmaceutical Research and Development, the United Kingdom (1960-1974).

[30] These variables are defined in comparable fashion to those for the U.S. case. See the Appendix for further details.

phenomenon of research depletion.[31] However, this line of reasoning is subject to at least two major qualifications. First, as noted above, the United Kingdom increased the scope of their regulatory controls over ethical drugs during the 1960s. Second, U.S. firms historically have been prominent in the U.K. market. Given this, it is plausible to expect that more stringent regulations in the United States after 1962 would have some negative "spillover" or "echo" effects on NCE introductions in the United Kingdom.

Relevant to this second point, we have plotted in Figure II the annual number of NCE introductions in the United Kingdom that were discovered in the United States.[32] This plot shows that U.S. discoveries introduced into the United Kingdom, exhibited a strong downward trend over the decade of the 1960s. Indeed this decline in U.S.-discovered introductions is a major factor underlying the downward trend in total U.K. introductions over this period. The observed pattern of U.S.-discovered NCEs in the United Kingdom is, therefore, quite consistent with the hypothesis of an echo effect from U.S. regulation postulated above.[33]

In order to minimize the biases associated with this interdependence phenomenon, we focus our analysis on domestically discovered NCE introductions. R & D productivity, the dependent variable of our analysis, is formulated as the number of NCE introductions originating in and developed by the pharmaceutical industry in each country relative to its R & D expenditures.

This procedure does not remove all of the bias associated with multinational interdependence, however. In particular, another problem arises from

[31] See his remarks as quoted at note 8 *supra*.

[32] The definition of a U.S.-discovered drug is the same one employed previously; that is, a drug discovered in a U.S. research laboratory, irrespective of the nationality of the laboratory ownership. See note 3 *supra*.

[33] It is interesting to note that the percentage of U.K. introductions accounted for by U.S. discoveries starts increasing during the seventies. In this regard, there are plausible reasons for expecting "echo" effects to be much greater in the short run (that is, the initial post-1962 period). This is because of the institutional procedures and strategies followed by U.S. firms in the preamendment period. In an earlier paper we found that, prior to 1962, most U.S.-discovered drugs were introduced in foreign markets, such as the United Kingdom, only after being introduced in the United States. Furthermore, many NCEs were initially manufactured here and exported abroad, in accordance with the product-life-cycle theory. Thus, at the time when regulatory conditions became more stringent in 1962, the rate of foreign introductions was quite directly tied to the level of U.S. introductions. In other words, foreign countries were generally treated as secondary markets by the U.S. firms.

As one might expect, the increased regulatory controls instituted in the United States after 1962 created strong incentives for firms to alter many of these traditional practices. Consistent with this viewpoint, we found a steady increase after 1962 in the percentage of U.S.-discovered drugs introduced in the United Kingdom before (or in lieu of) their introduction in the United States. Henry G. Grabowski & John M. Vernon, Innovation and Invention: Consumer Protection Regulation in Ethical Drugs, 67 Am. Econ. Rev. 359, tab. 2, at 363 (Papers & Proceedings, Feb. 1977). Nevertheless, this shift apparently took years to become fully effective—in part because of some significant legal barriers associated with the exporting of new drugs under review by the FDA. Henry G. Grabowski, *supra* note 9, at 51.

the participation of U.K. firms in the U.S. market. U.K. multinational firms obviously develop many of their products with the U.S. and other foreign markets in mind. As a consequence, increased costs of entry in the United States after 1962 would be expected to cause higher R & D costs and lower R & D productivity for many drugs discovered and developed wholly within the United Kingdom.

We hope this bias is second order in effect.[34] In any event, it will be similar in direction to the bias that comes from ignoring the effects of pre-1971 U.K. regulatory changes. In particular, our assumption that all changes of R & D productivity in the United Kingdom over the period 1960-1971, the control nation, are due to nonregulatory factors (and not due to increased regulation in the United Kingdom or the United States) will operate to produce an underestimation of U.S. regulatory effects.

In summary, a comparative international analysis does not provide an independent control like that of a laboratory experiment for two basic reasons. First, the regulatory environments in foreign countries like the United Kingdom have not remained completely fixed over time but have become more stringent in nature. Second, the drug industry has a significant multinational nature, so that increased regulatory controls in the United States would be expected to have some negative spillover effects on foreign country introductions and R & D activity. Although neither problem can be completely avoided, we hope to minimize the biases from spillover effects by focusing on R & D productivity (rather than total introductions) in each country. With regard to the biases which remain, we structure our analysis so that we obtain conservative estimates of regulatory effects. Thus, we wish to see whether a significant effect of regulation can be observed from our comparison of the United States and United Kingdom, even when the analysis is deliberately structured to produce an underestimate of regulatory effects.

C. *Simple Comparative Productivity Trends*

In this section, we present the basic comparative trends of the dependent variable for our analysis, R & D productivity. As discussed above, we use the term "productivity" to refer to the variable Baily defined as N_t/E_t, that is, the number of new chemical entities discovered and introduced in a country per effective R & D dollar. Following this, we present regression results,

[34] One reason for expecting this might be so is that our data suggest a much greater tendency for U.K. firms to license U.S. firms to develop and market drugs in the United States compared to the reverse situation involving U.S. introductions in the United Kingdom. One apparent reason for this is the unwillingness of the FDA historically to accept foreign trials as acceptable proof of safety and efficacy and its requirement that all applicable clinical trials be performed in the United States before considering a new drug application. (See Louis Lasagna & William M. Wardell, *supra* note 7, at 156.)

where the estimated U.K. time trend of productivity decline is used to represent the effect of all factors except regulation on U.S. productivity.

In Table 2, we show the productivity of R & D in the United States and the United Kingdom. Our initial calculations embody two of the strong assumptions made by Baily in his analysis. Specifically, 1) *all* R & D expenditures in each country are allocated to discovery of new NCEs[35] and 2) a five-year lag is assumed between R & D expenditures and the actual introduction of an NCE. These have been applied uniformly to the data for both countries. Since we are primarily interested at this point in the relative trend in R & D productivities of the two countries rather than the absolute value of R & D productivity at a point in time, these assumptions are less limiting than they might first appear. Furthermore, in our regression analysis in the next section, we relax the five-year lag assumption and allow for an increasing lag structure.

Because of U.K. data limitations, we were able to obtain productivities for only two years prior to 1962. However, for the later period we have measured productivity in five-year periods. These particular periods (1962-1966, 1966-1970, and 1970-1974) were selected because of the increased U.K. regulation which began in 1971. In addition, there has been a significant increase in R & D performance by U.S. firms in the United Kingdom and other countries in the 1970s, making the assumption of independence in the discovery process less tenable.[36]

TABLE 2
COMPARATIVE PRODUCTIVITY OF UNITED STATES AND UNITED KINGDOM IN DISCOVERED NCEs PER DOLLAR OF R & D INPUT

	United States		United Kingdom	
	Actual Value[a]	Index	Actual Value[b]	Index
1960-61	.232	594	.408	283
1962-66	.054	138	.232	160
1966-70	.039	100	.144	100
1970-74	.029	74	.061	42

Sources: See Appendix.

Notes:

[a] Number of NCEs discovered and introduced in the United States per R & D input (R & D is measured in millions of constant 1963 dollars).

[b] Number of NCEs discovered and introduced in the United Kingdom per R & D input. (U.K. data measured in millions of constant 1963 dollars where pounds are converted to dollar basis here at exchange rate of $2.80/pound).

[35] David Schwartzman, The Expected Return from Pharmaceutical Research 26-28 (Am. Enterprise Inst. 1975), has estimated that approximately 50% of the U.S. industry's ethical drug R & D expenditures over the period 1961-1967 were for the discovery and development of new NCEs as opposed to the development of other drug products (combinations, new dosage forms, and so forth). Thus, the assumption that all R & D is for new NCEs tends to somewhat understate R & D productivity in absolute terms (for both countries).

[36] See Henry G. Grabowski, *supra* note 9, at 44-48, for an analysis of the amount of R & D activity performed abroad by U.S. firms in recent years.

The productivities calculated in Table 2 should ideally be adjusted for any systematic differences in the quality of NCE introductions discovered in the United States and the United Kingdom. Teeling-Smith[37] has performed an analysis of the relative quality of discoveries in each country on all NCEs for which the first worldwide introduction occurred between 1958 and 1970. He found that U.S. discoveries for this period on average achieved a somewhat higher rating in terms of a quality index based on worldwide sales but a roughly comparable rating for a quality index based on medical importance (as evaluated by U.K. medical experts).[38] He concluded that a modest adjustment of the raw productivity calculation is warranted in comparing the two countries because of the higher overall quality of NCEs discovered in the United States. His findings in this regard are therefore consistent with somewhat higher (unadjusted) productivity for the United Kingdom in Table 2 for the initial period, 1960-61. Of course, this could also reflect differences in market structures, pre-1962 regulatory environment, and so forth.

Since our primary interest here is in the relative trends in productivity over time, we have included in Table 2 an index of productivities for each country, with productivity in 1966-1970 arbitrarily taken as 100.

The data presented in Table 2 clearly show that there has been a significant decline in the R & D productivities for the two countries over the postamendment period. However, perhaps the most interesting result is the much stronger *relative* decline in R & D productivity that the United States experienced in the decade after 1962. In particular, there is an approximate sixfold productivity decline in the United States and threefold decline in the United Kingdom between 1960-61 and 1966-70. Hence, over this period in which the United States shifted to a much more stringent regulatory environment than the United Kingdom, it also experienced a much more rapid decline in R & D productivity.

We should also note the steeper decline in productivity in the United Kingdom compared to the United States between 1966-70 and 1970-74. A plausible explanation for this phenomenon might be the onset of tighter regulation in the United Kingdom beginning in 1971.

Finally, the decline in the United Kingdom between 1960 and 1971 exhibited a much more steady trendlike character than in the United States. This is reflected in the data in Table 2 by the much more gradual rate of decline in R & D productivities in the United Kingdom over the successive five-year periods 1962-1966 and 1966-1970 than for the United States. When we estimated a time series regression of log N_t/E_t on time for the United King-

[37] G. Teeling-Smith, *supra* note 27.

[38] See *id.* In particular, Teeling-Smith found the weighted average market performance for U.S. compounds to be 2.8 million, while for the U.K. the average was 2.3 million.

dom over this period, we obtained a very good fit with an estimated annual rate of decline of 15 per cent. When alternative starting dates of 1961 and 1962 were used, the estimated rates of decline were 16 per cent and 15 per cent, respectively. Moreover, as noted earlier, the addition of an intercept dummy for 1962 or 1963 yielded statistically insignificant results, in sharp contrast to similarly estimated equations for the United States.[39]

Although these comparisons of simple R & D productivities are hardly definitive, they do suggest some important differences in the observed shifts in R & D productivities for these two countries. In the next section, we report the results of an econometric analysis in which we incorporate a measure of nonregulatory factors based on U.K. data into a production function model of the Baily type.

D. *A Regression Analysis of U.S. R & D Productivity*

In Part I (C), we reestimated Baily's model on U.S. data for the entire 1954-1974 period and found that his measure for depletion (that is, a moving average of past total introductions) became statistically insignificant. In this section, we analyze a similar production function model but make a number of significant changes in the basic functional specification.

i) *Controlling for Nonregulatory Effects Using U.K. Data.* The initial specification that we consider is:

$$\log [N_t/E_t] = a_0 + a_1 D_t + a_2 T_{pre60} + a_{UK} T_{post60}, \qquad (2)$$

where N_t = number of NCEs introduced and discovered by U.S. firms in year t

E_t = average industry-deflated R & D expenditures for ethical drugs in the United States in years t–4, t–5, and t–6 (it is assumed there is a fixed five-year lag from R & D outlays to introduction)

D_t = a zero-one dummy variable representing the effect of regulation (it equals 0 through 1961 and 1 afterward)

T_{pre60} = time trend representing 1954-1960 period (equals t from 1954 to 1960 and 7 thereafter, where t = 1 in 1954, 2 in 1955, and so on; see Appendix for details)

T_{post60} = time trend representing 1960-1974 period (equals 0 from 1954 to 1960 and $t - 7$ in 1961 and thereafter, where t = 1 in 1954, 2 in 1955, and so forth. See Appendix for details).

[39] See in particular the results presented in note 29 *supra* on this point.

In this specification, we estimate the effects of nonregulatory factors using a time trend calculated from U.K. R & D productivity data. In particular, we assume that in the absence of regulatory differences, R & D productivity in the United States would decline at an identical percentage rate as that for the United Kingdom. Under this assumption, the annual rate of decline of R & D productivity for the United Kingdom provides an external estimate of the impact of the nonregulatory factors for the United States.

In implementing this approach in terms of equation (2), the coefficient on the time trend variable after 1960 is restricted to equal the estimated decline in U.K. productivity after 1960. For the period before 1960, for which no U.K. productivity data are available, we use an unrestricted time trend to control for nonregulatory factors. The effects of the 1962 amendments are represented in this specification by the dummy shift variable D_t that takes on the value 1 after 1962 and 0 before.

Of course, the estimated rate of R & D productivity decline in the United Kingdom probably includes some negative effects from increased regulation in the United Kingdom as well as some "echo" effects for the United Kingdom of increased U.S. regulation. As argued above, we believe these echo effects are minimal since we are analyzing discoveries of U.K. origin rather than total introductions, but some effect is probably unavoidable. However, by attributing *all* of the decline to factors other than regulation, we will, if anything, obtain a conservative estimate of the impact of regulation.

In addition, the functional specification given by equation (2) retains a number of strong assumptions made by Baily as discussed in Section I (C) above. In the subsequent analysis, we will relax many of these assumptions.

The first step in estimating equation (2) is to estimate the annual rate of R & D productivity decline in the United Kingdom for the period 1960 to 1970. As noted earlier, least squares regression of the logarithm of N_t/E_t on time for this period yields an annual rate of decline equal to -0.15.[40] Restricting the coefficient on the *post*-60 trend variable to equal this value, we then estimate the other coefficients in equation (2) on U.S. data over the period 1954 to 1974. This yields the equation.

$$\log [N_t/E_t] = \underset{}{-0.49} - \underset{(1.72)}{0.85\, D_t} - \underset{(3.85)}{0.10\, T_{pre60}} - \underset{(1.71)}{0.15\, T_{post60}} \qquad \underset{\text{(restr.)}}{} \qquad (2')$$

$$R^2 = 0.92 \qquad F = 110.72 \qquad D.W. = 1.89.$$

[40] The least squares regression equation estimated for 1960 to 1970 in the United Kingdom was:

$$\text{Log}\left(\frac{N_t}{E_t}\right) = \underset{(4.00)}{1.39} - \underset{(5.43)}{.15\, T}$$

$$R^2 = .68 \qquad \rho = -.52 \qquad F = 17.22 \qquad DW = 2.44.$$

In effect, the restriction imposes a significantly faster annual rate of R & D productivity decline after 1960 compared to the estimated pre-1960 rate of 0.10. Furthermore, if one estimates equation (2′) without any restrictions on the trend variables, the least squares estimate on the post-1960 time trend variable is −.092, approximately the same as the estimated value on the pre-1960 trend variable. Thus, the restriction on the post-1960 time trend in equation (2′) clearly operates to amplify the implied effects of nonregulatory factors compared with the unrestricted situation.

Turning now to our main point of interest, equation (2′) further indicates that the regulatory shift variable D_t has a negative and statistically significant relation with R & D productivity. The estimated value of the D_t coefficient, −.85, implies that the 1962 amendments increased the average cost of a new NCE by a factor of 2.3. This is similar in magnitude to the rough calculations that we made on the basis of the productivity indices in Table 2.

The functional specification given by equation (2′) of course still retains a number of strong assumptions. In the analysis which follows, we relax a number of these assumptions in order to test the sensitivity of these results.

ii) *Alternative Functional Specifications.* We analyzed a number of alternative functional specifications to the log-linear formulation given by equation (2′). The best-fitting equation turned out to be the specification where the dependent and independent variables are all expressed in logarithmic units.[41] This formulation is presented as equation (3.1) in Table 3. It apparently results in an improvement in explanatory power over the log-linear case because it allows for a diminishing rate of productivity decline over time, rather than the constant rate implied in equation (2). However, aside from this difference, there is little change from the log-linear formulation. Indeed, the estimated coefficient on the regulatory shift variable, −.86, is virtually the same as before.

All the formulations analyzed to this point assume constant returns to scale between NCE introductions and past R & D expenditures. This assumption allows us to formulate our dependent variable as R & D productivity, N/E, and facilitates the econometric estimation of the model. As a check on the reasonableness of this assumption, we reestimated equation (3.1) (and the other variants of this model discussed below) with the inclusion of $\ln E_t$ on the right-hand side as another independent variable. The coefficients of $\ln E_t$ were never significantly different from zero and the estimated

[41] In this case, the restriction was based on the following equation estimated from U.K. data for the period 1960.

$$\operatorname{Log}\left(\frac{N_t}{E_t}\right) = \underset{}{3.89} - \underset{(4.94)}{1.76} \log T \quad \underset{(5.53)}{}$$

$$R^2 = .69 \qquad \rho = -.53 \qquad F = 17.65 \qquad DW = 2.52.$$

TABLE 3

REGRESSIONS USING LOG-LOG SPECIFICATION OF PRODUCTIVITY ON REGULATION AND TIME VARIABLES, WHERE COEFFICIENT OF LT_{post60} IS RESTRICTED TO EQUAL ESTIMATED TREND IN UNITED KINGDOM

Eq. No.	Dependent	Int.	D	LS	LT_{pre60}	LT_{post60}	R^2/F	DW	Period
				A. Fixed Lag Case					
(3.1)	Log (N/E)	−.55	−.86		−.28	−1.76	.94/147.31	2.44	1954-1974
		(2.21)	(4.90)		(1.67)	(restr.)			
(3.2)	Log (N/E)	.48		−.46	−.50	−1.76	.90/85.13	1.74	1954-1974
		(1.20)		(2.70)	(2.40)	(restr.)			
				B. Increasing Lag Case					
(3.3)	Log (N/V)	−.65	−.77		−.35	−1.21	.91/102.48	2.77	1951-1974
		(2.89)	(4.99)		(2.73)	(restr.)			
(3.4)	Log (N/V)	.35		−.45	−.49	−1.21	.86/64.45	2.13	1951-1954
		(1.04)		(3.08)	(3.25)	(restr.)			

Notes:

(1) t-statistics are given in parentheses.

(2) N = number of NCEs discovered and introduced by U.S. firms in year t.

(3) E = average deflated R & D expenditures in U.S. in years $(t-4)$, $(t-5)$, and $(t-6)$.

(4) V = "effective" R & D expenditures in year t assuming an increasing mean lag between R & D expenditures and NCE introduction (for details of construction, see Appendix).

(5) D = zero − one variable representing effect of regulation ($D = 0$ in 1954-1961 period and unity thereafter).

(6) LS = log of the continuous regulatory stringency variable S (see Appendix for details).

(7) LT_{pre60} = log of t from 1954 to 1960 and log of 7 in 1960 and thereafter, where $t = 1$ in 1954, 2 in 1955, etc. (see Appendix for further explanation).

(8) LT_{post60} = 0 from 1954 to 1960 and log of $(t/7)$ in 1961 and thereafter, where $t = 1$ in 1954, 2 in 1955, etc. (see Appendix for further explanation).

(9) In the increasing lag case, the definitions for the time variables were adjusted for the longer data period by setting $t = 1$ in 1951, 2 in 1952, and so forth.

coefficients on the other variables remained quite stable.[42] Hence, the constant-returns-to-scale assumption seems warranted.

We also tested the significance of the restriction imposed on the post-1960 trend variable for each specification in Table 3 by computing the appropriate F-statistic. Using the Wallace criterion,[43] the restriction could not be rejected at the 0.05 confidence level (critical values of F are tabulated in Goodnight and Wallace).[44]

[42] The estimated coefficients on ln E were positive in each case, but generally had t-statistics less than one in value.

[43] T. D. Wallace, Weaker Criteria and Tests for Linear Restrictions in Regression, 40 Econometrica 689 (1972).

[44] James Goodnight & T. D. Wallace, Operational Techniques and Tables for Making Weak MSE Tests for Restrictions in Regressions, 40 Econometrica 699 (1972). The computed F-statistics for the equations in Table 3 ranged from 0.10 to 1.45, all of which prevent rejection of the restriction at standard levels of significance.

In a strict sense, the estimated trend of U.K. depletion is not exact, but rather is an unbiased estimate of the trend which possesses substantial variance. If estimates of both mean and variance for coefficients of time trend variables are taken from the United Kingdom, they may be used in the method of J. Durbin, A Note on Regression when There Is Extraneous Informa-

iii) *Regulatory Stringency.* In our earlier discussion, we observed that the use of the zero-one dummy variable D_t to represent the effects of the 1962 amendments embodies a rather strong assumption. That is, it imposes the same shift factor on the entire postamendment period rather than a more plausible differential effect over time. To attempt to overcome this problem, we substitute a continuous proxy variable of regulatory stringency S_t for the shift variable D_t. In particular, our measure of S_t is the mean FDA approval time for a new NCE in each year (that is, the estimated time elapsing between the initial submission of a new drug application (NDA) and its final approval by the FDA). The available data on this question, which is admittedly quite crude, suggests FDA approval time steadily increased from seven months in 1962 until reaching a plateau of twenty-seven months in the period after 1967 (see the Appendix for further details).

Equation (3.2) of Table 3 shows the results of employing S_t to measure regulatory stringency, once again using the logarithmic specification of the model. The S_t variable is statistically significant and has the expected negative sign. Moreover, the estimated value of the coefficient suggests a cumulative impact from regulation that is comparable in magnitude to that previously estimated. In particular, it implies that increased regulation has caused the average cost per NCE to be larger in the post-1967 period by a factor of 1.86 compared to the pre-1962 period.[45]

It should be kept in mind that this measure of regulatory stringency, by its very nature, only considers drugs that successfully gain FDA approval. Another element of regulatory stringency which influences R & D productivity is the attrition rate on drugs that are clinically tested in man but fail to become NCEs. As discussed above, the attrition rate on clinically tested drugs has also significantly increased in the post-1962 period.[46] Hence, the development of a more composite index of regulatory stringency would seem to be a useful direction for further research.

iv) *Increasing Lag.* Another strong assumption embodied in all the model formulations estimated to this point is that the variable E_t assumes a fixed five-year lag between R & D expenditures and NCE introductions. Although

tion About One of the Coefficients, 48 J. Am. Stat. A. 799 (1953), to restrict coefficients in regressions for the United States. Due to the large variance of U.K. estimates, such inexact restrictions tend to be very much closer to unrestricted equations than those of Table 3. In other words, the statistically best use of information from the United Kingdom results in estimates of regulatory impact which are much higher and estimates of depletion—et al. impact which are much lower than are presented in the text. Again, the most conservative approach is taken.

[45] This was computed by substituting into equation (3.2) values of S of 7 months in the pre-1962 period and 27 months in the post-1967 period.

[46] If this element of regulatory stringency had a more direct and immediate impact on R & D productivity than lengthening approval times, which is not implausible, this may help explain why the D_t shift variable performs slightly better than S in Table 3. This is a question on which further research seems warranted.

good data is not available, there is considerable evidence which suggests that the average lag has increased significantly over the period we are studying.[47] Using the best estimates we could obtain on the average lag in different time periods, as well as some linear extrapolations, we constructed a variable lag variant of the equations estimated above. While the details of this construction are given in the Appendix, the basic assumption is that the average lag between expenditures and NCE introduction increased from 2.5 to 8 years over this period in the United States and increased by a somewhat lesser amount in the United Kingdom.

Equations (3.3) and (3.4) in Table 3 present the estimates for this variable lag variant of the model.[48] Essentially, the results are qualitatively similar to those given in the top half of Table 3. The estimates for this increasing lag formulation do indicate moderately lower impacts for the regulatory variables.[49] This is what one would expect, since an increasing lag over time (compared with the fixed lag used previously) operates to reduce the size of the decline in our R & D productivity dependent variable. However, it also should be kept in mind that an increasing lag *by itself* has a negative effect on innovative output and social welfare. Since it is commonly held that regulation is a major cause of this lag, it is appropriate to regard the estimated coefficients on D_t and S_t in equations (3.3) and (3.4) as only partial measures of the negative effects of regulation on innovative output and productivity.

To review briefly, all of the variants of the model analyzed imply a statistically significant and quantitatively important impact of the 1962 amendments. In particular, making conservative assumptions throughout, the estimated coefficients imply that increased regulation caused average costs per NCE to rise by a factor of between 1.8 and 2.3 over the first decade following the amendments. This amounts to more than one-third of the total increase in average costs experienced during this period.

E. *Qualifications and Possible Extensions*

It should be borne in mind that our analysis focuses only on the direct effects of regulation on R & D productivity or the average cost of discovering and introducing a new NCE. To the extent that increased regulation in fact has significantly increased the cost of introducing a new NCE, as our analy-

[47] L. H. Sarett, *supra* note 5.

[48] Ideally, the lag lengths and weights should have been estimated along with other coefficients, but multicollinearity and the paucity of data prevent this approach. The shift to a 2.5-year lag for early years made it possible to start regression analysis in 1951.

[49] Compared to the top part of Table 3 (that is, the fixed lag case), the implied effect of regulation on average cost per NCE changes from 2.36 to 2.16 in the case of the regulatory shift variable D_t and from 1.86 to 1.83 for the regulatory stringency variable S_t.

sis indicates, it should also affect the equilibrium level of industry R & D expenditures. In an expanded analysis, the total effect of regulation on NCE introductions, N, could be estimated by combining its effect on R & D productivity (N/E) with its effect on industry R & D expenditures E. The estimation of such expanded models would seem to a fruitful direction for further research.[50]

It may be noted that in a related analysis, David Schwartzman[51] has estimated the rate of return to pharmaceutical industry R & D for NCEs introduced over the period 1966-1972. He found a 6.6 pre-tax rate of return on R & D for this period, significantly below the average return on manufacturing investment and down from a 22.8 per cent return on pharmaceutical R & D in the early 1960s. If his estimates are correct, it would suggest that a significant part of the adjustment in equilibrium R & D has yet to occur. This is clearly a question on which more research would seem warranted.

Another important direction for further research would be to perform a more disaggregate analysis of R & D productivity in the two countries. William Wardell, a clinical pharmacologist, has compared the availability and therapeutic quality of NCE introductions in the United States and the United Kingdom after 1962 for a select number of therapeutic classes. He found a "drug lag" in the introduction of therapeutically beneficial NCEs into the United States compared with the United Kingdom, a lag which varied significantly in intensity across particular therapeutic classes.[52] It also would seem useful to compare R & D productivity in the two countries disaggregated by therapeutic class. This would allow one to see whether significant differences do exist and, if so, whether these differences might be plausibly associated with regulatory differences.[53] In order to undertake such an analysis, however, the necessary R & D data would have to be obtained from individual firm questionnaires, since these data are not presently available from public sources.

[50] We experimented with some simple reduced-form models on R & D expenditures that included regulation as well as various other supply-and-demand side factors as explanatory variables. Formulation of these equations on the basis of an optimality model incorporating our production function equation and a demand function results in a quite complex lag structure between R & D and the different explanatory variables. Using some very simple lag structures as a first approximation, we generally obtained the expected sign on the explanatory variables; but they were frequently not statistically significant. If one had a greater data base than the annual time series observations available here, one could presumably estimate these equations in a more precise fashion.

[51] David Schwartzman, *supra* note 35, at 36.

[52] For a summary of this work see Louis Lasagna & William M. Wardell, *supra* note 7, Part II, at 51-123.

[53] For example, it is presumably much easier to prove efficacy for an antibiotic than for several other classes such as cardiovascular drug therapies. Wardell found a much greater drug lag in the latter case compared to the former one. It would be useful to see if such patterns also emerge in a comparison of R & D producturies.

IV. Summary and Conclusions

Innovation in the pharmaceutical industry has been subject to a number of adverse structural developments in recent years. There has been a sharp decline in the annual number of introductions of new chemical entities and rapid increases in costs and risks. We have reviewed these developments and listed five hypotheses that have been used to explain them: (1) increased regulation of the industry associated with the 1962 amendments to the Federal Food, Drug, and Cosmetic Act is the cause; (2) the decline is illusory since only ineffective NCEs have declined; (3) a depletion of research opportunities has taken place; (4) the thalidomide incident has made firms and physicians more cautious; and (5) costs have risen as a result of advances in the technology of safety testing.

In order to separate the effects of regulation from these other confounding factors, we developed an international comparative analysis of R & D productivity changes in the United States and the United Kingdom.

A principal finding that emerges from this international comparative analysis is that U.S. "productivity"—defined as the number of new chemical entities discovered and introduced in the United States per dollar of R & D expenditure—declined by about sixfold between 1960-61 and 1966-70. The corresponding decrease in the United Kingdom was about threefold. Clearly, some worldwide phenomenon, which might be labelled a "depletion of research opportunities"—but which probably also includes the effects of other factors such as the thalidomide incident and higher costs due to new developments in safety testing—seems to hold for pharmaceutical R & D. However, there is also strong support for the hypothesis that an additional factor has been at work in the U.S. industry.

We conclude that this additional factor, which has lowered U.S. productivity at a significantly more rapid rate, is the increased regulation resulting from the 1962 amendments. On the basis of the regression analysis presented in Section III, we estimate that the 1962 amendments have probably, at a minimum, doubled the cost of a new entity.

Our analysis also suggests that nonregulatory factors have an important aggregative effect on innovation, but does not allow us to say which factors in particular have been most important in this respect. Further research on this question would seem warranted.

APPENDIX

This appendix presents in summary form the sources and methods of computation for statistics used in the paper.

NCE Introductions and Discoveries

Data on new chemical entities and their years of introduction for both the United States and the United Kingdom were obtained from the publications of Paul de Haen.[54] In a very few cases, information on British introductory dates was supplemented by the work of William Wardell.[55] Biologicals and diagnostics were here deleted from data lists and analysis due to problems of data availability and reliability prior to 1966.

Information as to which of these NCEs were also discoveries by industry research laboratories was obtained for the United States from Paul de Haen,[56] for the United Kingdom in 1960-1970 from the National Economic Development Office,[57] and for the United Kingdom in 1970-1974 from, again, Paul de Haen.[58] An NCE was regarded as discovered in a particular country if the research laboratory producing the entity was located in that country, irrespective of the nationality of laboratory ownership. Thus the discoveries of Pfizer in the United Kingdom are credited to Britain while those of Hoffmann-La Roche in the United States are considered as American. It should be recognized that the discoveries of NCEs are denoted by year of introduction in either the United States or the United Kingdom (depending on origin) rather than first year of introduction on a worldwide basis (should these dates differ).

R & D Expenditures

Expenditures for research and development are here considered as those domestic outlays by the pharmaceutical industry for discovery of humanly usable ethical drugs. In the United States, data were obtained from publications of the Pharmaceutical Manufacturers Association (PMA)[59] for worldwide human R & D expenditures, 1948-1974, of member firms. However, the breakdown of domestic versus foreign

[54] Paul de Haen, Compilation of New Drugs, 33 Am. Professional Pharmacist 25-62 (Nov. 1967); *id.*, 7 New Drug Analysis USA, 1966-1970 (1971); *id.*, 10 New Drug Analysis USA 1969-1973 (1974); *id.*, New Products Parade (20th ed., mimeographed, Feb. 1975); *id.*, New Single Drugs Marketed in England, France, Germany, and Italy 1960 to 1965 (mimeographed, Feb. 1973); *id.*, New Single Drugs Marketed in England, France, Germany, and Italy 1966, (mimeographed, Oct. 1973); *id.*, 1 New Drug Analysis Europe, 1967-1971 (1972); *id.*, 4 New Drug Analysis Europe, 1970-1974 (1975).

[55] W. M. Wardell, Introduction of New Therapeutic Drugs in the United States and Great Britain: An International Comparison, 14 Clinical Pharmacology & Therapeutics. 773-90 (1973).

[56] Paul de Haen, Compilation of New Drugs, 33 Am. Professional Pharmacist 25-62 (Nov. 1967); *id.*, 7 New Drug Analysis USA, 1966-1970 (1971); *id.*, 10 New Drug Analysis USA, 1969-1973, (1974); *id.*, New Products Parade (20th ed., mimeographed, Feb. 1975).

[57] National Economic Development Office, A List of 466 Pharmaceutical Compounds and Country of Discover (mimeographed, 1971) (prepared for NEDO by the Centre for the Study of Industrial Organization as part of the study, Innovative Activity in the Pharmaceutical Industry).

[58] Paul de Haen, 1 New Drug Analysis Europe, 1967-1971 (1972); *id.*, 4 New Drug Analysis Europe, 1970-1974 (1975).

[59] Pharmaceutical Manufacturers Association, Annual Survey Report (various years); *id.*, Office of Econ. Research, Prescription Drug Industry Factbook (1967).

expenditures in this total was available only for 1960-1974, from the same sources. By fitting an exponential trend for foreign R & D expenditures of PMA member firms against time, 1960-1974, estimates of this parameter were obtained for earlier years. Subtraction of these estimates from the worldwide total gave the data used in the text.

R & D data for the United Kingdom for 1954-1966 and 1973 were taken from releases of the Association of the British Pharmaceutical Industry.[60] For 1954 to 1965, the data aggregated human and veterinary research expenditures. These statistics were multiplied by 86.1 per cent (the 1966 value) to obtain estimates of expenditures for purely human research. For the years 1966 to 1974 an exponential trend on time was fitted to obtain R & D estimates for intervening years.

R & D estimates for both industries were deflated by the gross national product deflator to constant (1958) dollars for the United States[61] and to constant (1958) pounds for the United Kingdom.[62] Statistics for deflated expenditures on R & D as well as introductions and discoveries of NCEs are plotted in Figures I and II of the text.

Pharmaceutical Sales

Data on U.S. sales of ethical drugs were obtained from the publications of a marketing research firm, Intercontinental Medical Statistics.[63] These data were based on a projection from a 1,000 drug store sample to the population of all U.S. drug stores, and on a sample of about 10 per cent of total hospital beds. Sales directly to other institutions, such as to the U.S. government are here excluded, but they account for less than 20 per cent of U.S. ethical drug sales.

FDA Stringency

Estimates of the mean time in months to FDA approval of NCEs introduced in the United States were taken from an unpublished dissertation of Joseph M. Jadlow.[64] Jadlow obtained his estimates through private communication with the FDA. The figures used in the text extrapolate from Jadlow's and are as follows:

1954-1961	7.0 months
1962	9.3 months
1963	11.3 months
1964	14.0 months
1965	19.0 months
1966	24.0 months
1967-1974	27.0 months

[60] Association of the British Pharmaceutical Industry, Annual Report 1973-1974, (1974); *id.*, Pharmaceutical Research and Development Survey (mimeographed, Jan. 17, 1975).

[61] Economic Report of the President, Together with the Annual Report of the Council of Economic Advisors (1975).

[62] Central Statistics Office, Annual Abstract of Statistics (London, various years).

[63] Intercontinental Medical Statistics, Pharmaceutical Market—Hospitals (various years); *id.*, Pharmaceutical Market—Drugstores (various years).

[64] J. M. Jadlow, Jr., The Economic Effects of the 1962 Drug Amendments 174 (1970) (unpublished Ph.D. dissertation, University of Virginia).

These values are defined as the variable S, the logarithm of which is used in equations (3.2) and (3.4) of Table 3.

Lags for Effective R & D Expenditures

Estimates of development times for NCEs were interpolated from figures offered by Dr. Lewis Sarett.[65] Addition to these development times of the regulatory approval times given above yields the following estimates of total lag times, from first expenditure to introduction:

1954-1958	2.5 years
1959	3.0 years
1960	3.25 years
1961	3.5 years
1962	4.0 years
1963	4.65 years
1964	5.25 years
1965	5.8 years
1966	6.4 years
1967	7 years
1968	7.3 years
1969	7.65 years
1970-1974	8 years

R & D expenditures in a given year become effective over a three-year period centered around the (mean) total development period. For example, expenditures in 1967 are seen as effective in 1973, 1974, and 1975 at the rate of one-third of original 1967 expenditures. Total effective expenditures are obtained by summing over all expenditure portions which become effective in the given year and are defined as the variable V in Table 3. While admittedly stylized, this lag system appears to capture the essence of the process at issue. Further, alternative lag structures based on the above mean lag estimates, as well as minor alterations of the mean lag estimates themselves, yielded qualitatively similar results in all cases.

It should also be noted that in estimating the U.K. trend for the restriction in the increasing lag case, an increasing development period ranging from two to five years was assumed.

Mechanics of Estimation

The specification assumed for equation (2) in the text can be written as:

$$\log(N/E) = a_0 + a_1 D + a_2[(1 - X)\,t + 7X] + a_3 X\,(t - 7),$$

where (1) a_3 is restricted to equal U.K. trend
(2) t is 1 in 1954, 2 in 1955, . . .
(3) $X = 0$ from 1954 to 1959 and unity thereafter.

Hence, the variable T_{pre60} in equation (2) is the multiplier of a_2 above and T_{post60} is the

[65] L. H. Sarett, *supra* note 5.

multiplier of a_3. The reason for the rather complex definitions of these two time trend variables is to ensure that the two time trend segments join properly in 1960. Thus, a_2 is the rate of decline of N/E from 1954 to 1960 and a_3 is the rate of decline thereafter.

Similarly, the specification of the log-log version of the above equation, equation (3.1) in Table 3, can be written in terms of t and X as follows:

$$\log(N/E) = b_0 + b_1 D + b_2[(1 - X)\log t + X\log 7] + b_3(X\log t - X\log 7),$$

where b_3 is restricted to equal U.K. trend.

Thus, as above, the variable LT_{pre60} in Table 3 is the multiplier of b_2 above and LT_{post60} is the multiplier of b_3.