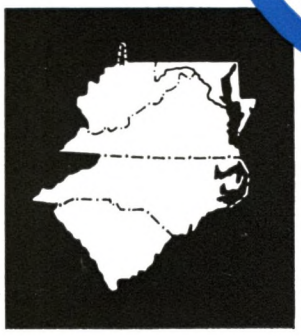
FEDERAL RESERVE BANK OF RICHMOND

MONTHLY REVIEW

The Dismal Science Revisited

Bank Affiliates and Their Regulation: Part I



Volume 59
Number 3

MARCH 1973

THE DISMAL SCIENCE REVISITED

For the past several years economic growth has been the focus of heated debate. Controversy continues to rage over such growth-related issues as the population explosion, the green revolution, the energy crisis, environmental pollution, ecological destruction, and natural resource exhaustion. On one side of the debate are the pessimists who, ever fearful of the harsh constraints of nature, perceive the imminent demise of growth and warn darkly of approaching calamity and doom. At the other end of the spectrum are the optimists who are skeptical of apocalyptic prophecies and believe that the limits to growth are not yet in sight.

As seen by the pessimists, the basic problem is simply the impossibility of continued exponential (constant percentage) growth in a world of limited resources. In an exponentially-expanding system, each year's growth adds a larger absolute increase because it is applied to a bigger base; consequently, growth accelerates rapidly like compound interest. Pessimists note that such is the awesome power of compound interest that quantities growing at a geometric rate will ultimately surpass the largest finite magnitude. Eventually, therefore, resource constraints must become binding. What really alarms the pessimists, however, is the swiftness with which exponential growth paths approach a fixed limit. If the doubling time of population is 20 years, it may take centuries to reach a point half way to the ceiling; but it only takes 20 more years to go from the halfway point to the ceiling. Pessimists believe that current growth paths—of population, pollution, and output—are steepening and are on a catastrophic collision course with ceilings imposed by limited and dwindling stocks of extractive and environmental resources. When the paths hit the ceilings, their direction will be reversed; thereafter, population and output will plunge downward as famine and resource depletion exact a fearsome toll. In support of these dire forecasts, pessimists cite computer simulations showing doomsday a mere two generations away.

The optimists, in response, point out that pessimists overlook such offsets to scarcity as technological progress, increased knowledge, and resource substitution. These scarcity-ameliorating factors, it is claimed, may permit the ceilings to rise at rates rivaling those of population and output growth. Optimists also contend that the doomsday computer models do not take adequate account of certain feedback mechanisms that induce stabilizing alterations in human behavior patterns. Optimists believe that such mechanisms would operate to slow or stop growth if resource ceilings were approached. In other words, adjustment mechanisms would transform accelerating, explosive growth paths into decelerating, convergent ones.

Speculation on the subject of implacable constraints to growth is by no means a new pastime. Gloomy prognostications of impending natural resource scarcity have a long tradition, extending back at least to the early nineteenth century when Thomas Malthus, David Ricardo, and other economists of the British classical school prophesied that unlimited procreation combined with limited land would bring diminishing returns, bare subsistence incomes, and the eventual cessation of growth. In fact, the pessimistic doctrine of resource scarcity formed the core and central theme of classical political economy, thereby providing economics with its reputation as the "dismal science." Later, at the beginning of the twentieth century, the dismal theme of scarcity was enunciated again, this time by conservationists who stressed the resource-depleting and resource-exhausting effects of economic growth.

The similarity between these earlier writings and contemporary discussion is not always fully appreciated. In many cases the modern doomsday prophets are merely echoing doctrines stated by classical economists and conservationists. Virtually all of the elements of the current doomsday models—exponential population growth, a shrinking or static resource base, the concept of the stationary state—appeared in earlier writings. Moreover, the reversal of economists' views on the issue of resource limitation has not been generally recognized. The "dismal science" label is no longer properly descriptive of the outlook of many modern economists, who tend to be somewhat more sanguine than their classical predecessors.

It is not modern economists, but rather those conservationists, ecologists, and biologists that continue to write in the classical economic tradition, who are the current representatives of the pessimistic interpretation of the economic problem.

How did this switch occur? What influences induced economists to become optimists and scientists pessimists? How was the problem of resource scarcity originally formulated and analyzed? In what forms do the older doctrines survive in current debates? Has history provided any empirical tests of the older doctrines? Does the spectre of resource limitation, when viewed in historical perspective, become more terrifying or less? These issues are examined in this article, which outlines the evolution of the resource limitation doctrine from the writings of the classical economists to the current doomsday debate.

I. CLASSICAL ECONOMISTS AND THE PROBLEM OF NATURAL RESOURCE SCARCITY

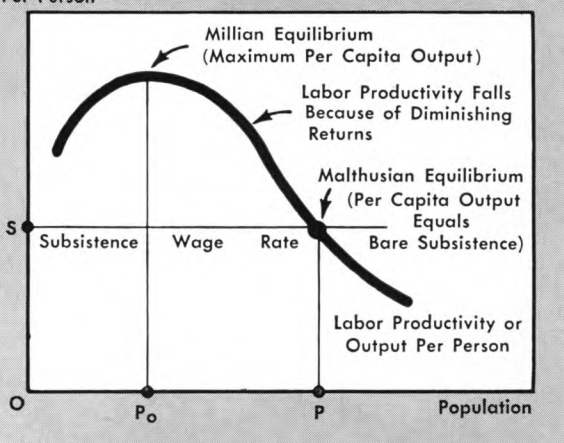
The first rigorous analytical statement of the problem of growing natural resource scarcity appears in the early nineteenth century writings of economists of the British classical school. Classical economists feared that low agricultural productivity eventually might threaten to obstruct British industrial development. The agricultural productivity problem was thought to stem from the growing disparity between a fixed supply of land and a rapidly burgeoning population. Classical economists interpreted the land-man divergence within the framework of their analytical model, which they used in predicting the effects of limited natural resources on the pace and character of long-run economic growth, as well as on the behavior of the relative shares in the national income.

Malthusian Population Growth The chief expositors of the classical model were Thomas Malthus, David Ricardo, and John Stuart Mill. Malthus is best known for his celebrated population doctrine, which states that man's propensity for procreation far exceeds his capacity to expand the food supply. According to the Malthusian theory, population, if unchecked, tends to multiply at an exponential or constant percentage rate, with the annual absolute increment to the population becoming successively greater with time. Malthus himself estimated this population growth rate to be approximately 3 percent, close to the biologically maximum sustainable rate that would result in a doubling of the population every 25 years. Owing to the fixed supply of agricultural land, however, the food supply cannot be increased at the same rate; or as Malthus expressed it, population expands geometrically, food only arithmetically.

The Law of Diminishing Returns Malthus and other classical economists employed the *law of diminishing returns* to explain why the food supply could not expand in equal proportion with the number of laborers working the soil. According to the law of diminishing returns, the application of larger and larger amounts of labor to fixed land eventually results in declining average labor productivity, i.e., output per worker. With fixed land and increasing population, the land/labor ratio must fall. Since each individual worker will have less and less land to work with as more and more workers crowd on the fixed land, average labor productivity, or crop yield per worker, must fall. Multiplication simultaneously creates new stomachs and new hands. But, alas, while the food requirements of each stomach remain constant, the food production of additional pairs of hands continually declines. Consequently, the inability of food production to expand as rapidly as the biological growth of population serves as the ultimate check to population growth. Specifically, the Malthusian limits to population growth occur when diminishing returns to labor working with scarce land bring crop yields down to the minimum level of subsistence. (See Chart 1.)

Chart 1

MALTHUSIAN vs. MILLIAN POPULATION EQUILIBRIA



Beyond P_o, population growth is accompanied by diminishing returns, causing labor productivity to fall. Malthus thought population would stabilize at P, the maximum size sustainable at bare subsistence. Mill was hopeful that voluntary birth control would limit population to its optimal size, P_o, where wages were highest.

The Ricardian Growth Model Malthus stressed diminishing returns resulting from the crowding of labor on fixed land of constant quality. David Ricardo emphasized another source of diminishing returns, namely, the cultivation of agricultural land of progressively inferior quality. Ricardo's chief contribution, however, was to combine the Malthusian population doctrine, the law of diminishing returns, and a theory of capital accumulation into a comprehensive interpretation of the effects of natural resource scarcity on economic growth and income distribution.

In the Ricardian model, the process of growth is seen as an inexorable movement of the economy toward its long-run, zero-growth stationary equilibrium. Along the path to the stationary state, diminishing returns to labor and capital employed on the scarce land act to raise land rents, to lower wages and profits, and finally to bring growth to a halt. Ricardo's model concentrates almost exclusively on the land-using agricultural sector. Ricardo's neglect of the manufacturing sector stemmed from his belief that the growth-inhibiting influences operative in food production would eventually swamp the growth-stimulating influences operative in industrial production. Like the other classical economists, Ricardo thought that the non-land using manufacturing sector would experience constant or even increasing returns to labor-capital inputs, yet this constant or rising productivity in manufacturing would be more than offset by diminishing returns in land-intensive agriculture. Similarly, technological progress, a force that tends to counteract or offset diminishing returns, was disregarded in the Ricardian model as being of relatively minor importance, i.e., historically diminishing returns were assumed to dominate technological advance.

The Ricardian analysis begins with a hypothetical economy in an early stage of development. Population is small in comparison with the available land. Because land is at first relatively plentiful, rents are low; but wage rates, profits, and the rate of capital accumulation are all high. Capital accumulation, the chief engine of growth in the Ricardian model, serves to promote production. It also stimulates the demand for labor, thereby bidding the market wage rate temporarily above the subsistence level. Via the Malthusian mechanism, the above-subsistence wage rate then induces population growth, which, in turn, necessitates a more intensive and extensive cultivation of the land. But at some stage the application of additional labor to the limited supply of land brings diminishing average returns, which become more pronounced as growth proceeds. Hence, average labor productivity continually declines as labor presses ever more densely on good land and spills over onto poor land. This development sequence, however, benefits the landowning class, which receives rents in proportion to the excess of productivity on the better parcels of land over the least productive acres in use. Extension of the margin of cultivation increases the premium or differential surplus return received on the better ("supramarginal") land. Thus, land rents rise; and the landowners' share of the national income expands as growth increases the demand for land.

The portion of output remaining after rents have been paid is distributed between workers and capitalists. As long as this residual output is in excess of the wage bill, there will be profits available to finance the capital accumulation that keeps the whole developmental sequence going. Over time, however, diminishing returns will act to close the profit gap between net (of rent) output and the subsistence wage bill, as shown on Chart 2. As profits are squeezed, capital accumulation slackens. Eventually, profits will vanish under the pressure of diminishing returns. At this point all growth stops, capital accumulation ceases, population expansion halts, and output stabilizes at a constant level. The economy has reached its zero-growth, long-run equilibrium—the classical stationary state.

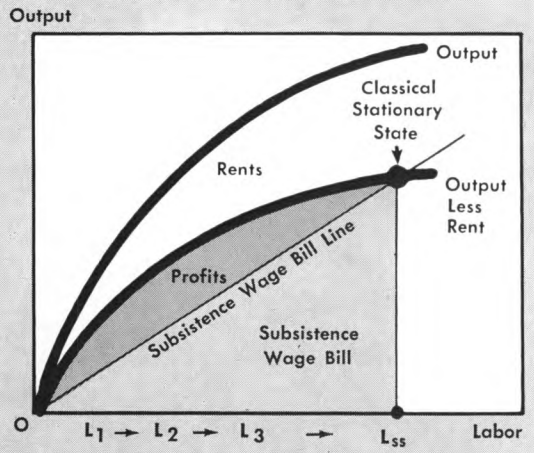
The Classical Stationary State What would conditions be like in the Ricardian stationary state? For one thing, there would be a marked inequality in the distribution of incomes. Rents would be high, wages low, and profits virtually non-existent. A few wealthy landowners would be enjoying a life of splendor and ease, while the mass of the population would be muddling along at subsistence levels of income. The total mass of poverty would indeed be great because the population size would be the maximum sustainable on the limited land, given the state of technology. All output in excess of the portion claimed by landowners would be absorbed just to support the working population at subsistence, i.e., a level of income sufficiently high to allow workers to replace themselves without increase or diminution.[1]

The dominant feature of the stationary state, however, would be the complete absence of change. There would be no technical progress, no innovation, no new products, no new occupations, and no advances

[1] Classical economists believed the subsistence wage to be somewhat higher than the biological or physiological minimum for survival. They acknowledged that the subsistence standard of living was in part psychologically determined by prevailing habits and customs. Nevertheless, they thought this long-run natural wage level to be substantially below what could be described as a "comfortable" standard of living.

Chart 2

RICARDIAN GROWTH MODEL



Profits fuel capital formation, which stimulates growth and raises the market rate of wages temporarily above the subsistence level. Above-subsistence wages induce procreation and thus expansion of the labor force (sequence $L_1 \rightarrow L_2 \rightarrow L_3$ etc.). But more labor applied to fixed land brings diminishing returns (shown by the declining slope of the output and output-less-rent curves). Diminishing returns raise land rents but lower profits and drive wages back to subsistence. Growth slows as profits are increasingly squeezed between diminishing returns and subsistence wages. When profits vanish, growth stops and the economy reaches the stationary state.

in knowledge. Population would remain forever constant in size, age composition, and skill. The stock of capital, too, would remain unchanged in both size and structure. Balance conditions would insure the eternal constancy of the capital and population stocks. Births would balance deaths. Production would balance consumption. Capital replacement would just balance capital depreciation. All activity would be devoted to stock maintenance and replenishment; none to growth.

In general, Malthus, Ricardo, and other economists of the British classical school were apprehensive about the advent of the stationary state. To them it offered only the chilling prospect of the mass of humanity forever doomed by the limited resource base to low wages. There would be no escape, they thought, from the stationary state. Being a stable equilibrium, it was like a trap. Any temporary departure from the subsistence equilibrium would automatically set in motion forces that would bring the system back to the equilibrium position again. If a small disturbance caused the population to deviate above or below the equilibrium size then, inevitably, starvation or procreation would bring it back to the stationary level. In short, there would be no exit from the stationary state, which was indeed a pessimistic conclusion. Small wonder that Malthus and Ricardo were described by Thomas Carlyle as "the dismal practitioners of the dismal science."

A Cheerful Stationary State? Of the classical economists, only John Stuart Mill was sanguine about the stationary state. But Mill's cheerful version of long-run equilibrium differed considerably from that of his gloomy contemporaries. Whereas the latter thought that stasis was an inevitable condition imposed by the constraints of a harsh and unyielding nature, Mill saw it as a possible outcome of workers' voluntary decisions to restrict births while resources were still abundant. Mill believed that there was some critical threshold level of per capita income, above which birth rates would decline as income increased. He was hopeful that capital accumulation and technical progress might raise and maintain wages above the critical level before diminishing returns had gone too far. Then, workers, aspiring to protect or improve their newly-established standard of living, would exercise more prudence, foresight, and self-restraint in procreation. Hence, the advent of Mill's stationary state might occur with the population stabilizing at its *optimum* size consistent with peak per capita income rather than its *maximum* size consistent with the bare subsistence level. (See Chart 1.) Thereafter, continued voluntary restraint on propagation would insure that per capita income would never decline. Moreover, technological progress might provide for its continued increase. Unlike the other classicals, Mill recognized that technological progress could proceed independently to raise incomes even though capital and population were stationary. In other words, Mill's stationary state was not stationary with regard to technology.

Mill also had a far more sophisticated conception of natural resources than did the other classical economists. The latter viewed the natural resource base as consisting solely of *land*—an indestructible, inexhaustible (yet not reproducible) asset yielding an output in the form of physical product, i.e., agricultural commodities. Mill, however, called attention to the *open-space* dimension of land. Open space, he noted, is an exhaustible asset yielding an output in the form of *amenities*, e.g., personal solitude, scenic grandeur, etc. Mill's perceptive and original observation presaged recent analyses of environmental, as distinct from commodity, resources.

Mill's insight, however, was atypical of the classical school. Among his classical contemporaries, Mill was the exception. For the most part, classical economists ignored technical progress as an offset to

diminishing returns; slighted the possibility of widespread voluntary restraint of births; viewed the stationary state with foreboding; and saw the problem of natural resource scarcity solely as one of too little land relative to the population. In sum, the classicals' interpretation of resource scarcity was limited in scope and slanted toward pessimism. Their interpretation would have been even more pessimistic, however, had they viewed the resource base as a wasting, rather than a fixed, asset. But resource depletion was not stressed in their analysis. In fact, it was not until much later that other writers switched the emphasis to problems of resource depletion. The first *economist* to do so was William Stanley Jevons who in 1865 warned that incipient coal depletion and the consequent rising cost of coal extraction would shortly signal the end of British industrial supremacy. But Jevons was virtually the only major economist up to the twentieth century to examine fully the problem of resource exhaustion. The main thrust on the depletion issue was to come from conservationist writers, rather than from economists.

II. CONSERVATIONISTS AND THE PROBLEM OF NATURAL RESOURCE ATTRITION

Concern over natural resource scarcity re-emerged at the beginning of the twentieth century when American conservationists warned of excessive depletion and imminent exhaustion of U. S. stocks of extractive resources. Thus, to the classical spectre of diminishing returns, conservationists added the dismal prospect of a shrinking resource base.

Conservationist Tenets Several unique features characterized the conservationist position. First, early conservationists often harbored an anti-market bias. They suspected that a private property, free-market economic system inevitably promoted excessive rates of resource depletion. Specifically, conservationists claimed that the profit motive induces, and private ownership allows, the rapacious and profligate current exploitation of resources. Resource owners were criticized for selfishly disregarding the needs of the future in their impatience to maximize present profits. The market, too, was said to be short-sighted, sacrificing high-priority future uses for low-priority present ones. Thus, the time distribution of resource use was described as being strongly biased against the future; natural wealth, it was claimed, is squandered in the present at the cost of an impoverished posterity. Owing to their distrust of the price system, conservationists advocated resource management by government agencies. Such agencies, it was believed, would prevent current resource waste and would preserve resource stocks for the benefit of future generations.

Second, conservationists adhered to a *physical,* as contrasted with an *economic,* concept of resource wealth. According to the economic concept, natural resource wealth is measured by the present discounted value of the expected future stream of services to be yielded by the stock of virgin resources. Since society's wealth may be held in forms other than idle natural resources, the present values of these alternative forms also can be calculated to determine which types of assets are more socially productive, i.e., which will produce a larger flow of future services. Using the present value concept, economists have demonstrated that investments in education, research, technology, and reproducible capital equipment have been far more productive of future and present real income than the mere holding of idle natural resources would have been. Older conservationists, however, often overlooked this point. Their single-use concept of wealth blinded them from the realization that natural resources might be more valuable and productive if allocated to other uses. To conservationists, proper resource management involved the husbanding of physical stocks of particular resources rather than the maximization of present economic value. Conservationists recommended the compilation of elaborate inventories of reserves of physical resources. These inventories, it was assumed, would adequately represent the value of the resource legacy transmitted from the present to the future. Overlooked, however, was the likelihood that technical change might alter the economic value and the quality-ranking of an inventory's components. Generally, the possibilities of low long-run correlations between economic values and physical resource quantities were disregarded.

Third, conservationists stressed the concept of ecological interdependence. Long before modern ecologists, early conservationists were discussing environmental resources and the "fragility" of ecological equilibrium. Thus, the conservationists' concept of resource attrition included ecological destruction as well as mineral-resource depletion.

In stressing resource depletion, conservationists contributed the crucial closing link to the resource scarcity doctrine. The classical economists had emphasized diminishing returns stemming from a growing demand impinging on a fixed resource supply. To this, conservationists added the element of a contracting supply, thereby making the dismal doctrine even more dismal. Thereafter, resource scarcity was interpreted as a problem of demand expansion intensified by a shrinking supply.

Dominance of the Dismal View The classical prognosis went virtually unchallenged until the 1930's. It is true that in the late nineteenth century a few British economists, influenced perhaps by a complacent Victorian belief in evolutionary progress, admitted that technological progress might postpone for several generations the arrival of the classical stationary state. Yet even these economists left no doubt that they believed that the race between fecundity and technology would eventually be won by the former. Thus, the vision of the inevitable stationary state persisted, at least as a remote historical destiny if not as an imminent threat.

The classical outlook continued to influence thinking in the post-World War I decade. For example, in the 1920's, John Maynard Keynes was arguing that Malthusian and Ricardian pressures emanating from the underdeveloped countries were threatening to end economic progress in Britain. Keynes contended that diminishing returns operating in the primary product (food and raw material) producing sectors of Britain's major foreign suppliers were being transmitted to Britain via a secular deterioration in the British terms of trade. With diminishing returns abroad raising the price of primary-product imports relative to the price of British manufactured goods, more and more exports of the latter would have to be sacrificed to obtain a unit of food and raw material imports. Keynes feared that the rising real cost of primary-product inputs would check industrial progress and lower the standard of living in Britain.

Keynes's restatement in the 1920's of the Ricardian conclusion marked the close of an era in which the pessimistic resource scarcity doctrine had commanded almost universal allegiance (or at least perfunctory lip service) among economists. After reigning as part of the conventional wisdom for more than a century, the doctrine was to be severely challenged by events and research studies in the next four decades.

III. CHALLENGES TO THE RESOURCE SCARCITY DOCTRINE

Emergence of Skepticism Belief in the classical-conservationist outlook suffered substantial erosion in the 1930's when a sharp drop in birth rates seemed to presage future *de*population rather than Malthusian overpopulation. Moreover, the classical theory of the stationary state was powerless in diagnosing the chronic secular stagnation, thought in some quarters to be foreshadowed by the Great Depression of the 1930's. According to the classical theory, long-run stagnation is imposed by insufficient resources relative to demand. Yet, according to some analysts, the industrial stagnation that threatened mature capitalistic economic systems in the 1930's appeared to stem from a persistent deficiency of aggregate demand relative to resources. In sum, events in the 1930's weakened the credibility of the classical-conservationist doctrine. But the real blow came in the 1950's and 1960's in the form of empirical and analytical findings that refuted classical-conservationist predictions.

Empirical Findings The classical-conservationist resource scarcity doctrine, it will be recalled, predicted that diminishing returns and the growing relative scarcity of land and natural resources would: (1) lower the productivity of labor, thus driving real wages and income per capita to subsistence levels; (2) raise the share of natural resource owners in the national income while simultaneously squeezing the profit share until the incentive to save and accumulate capital had vanished; (3) raise labor-capital cost per unit of extractive output absolutely as well as relative to the cost of nonextractive output; and (4) slow and ultimately stop output and capital growth. Moreover, the classical-conservationist assumption of constant technology denied the possibility that natural resource scarcity might be offset by resource-saving innovations and by the increasing feasibility of substituting capital for land in productive processes.

These forecasts have not stood up well when confronted with empirical data. Studies of long-term economic growth show that the predicted events have not yet materialized—at least not in the U. S. and other developed countries of the West. Instead, almost the opposite of what had been predicted has actually happened. Productivity, real wage rates, and real income per capita have registered manifold increases since the nineteenth century. Long-term percentage growth rates of output and capital have shown little or no tendency to fall. There has been no long-run decline in either the rate of profit on capital or the profit share of national income. Instead, contrary to Ricardian predictions, the income share of natural resource owners has fallen to one-third or less of its nineteenth century level, indicating the declining relative importance of natural resources in the production process. Natural resource inputs per unit of output have fallen greatly, as has also the percentage share of the GNP consisting of resource products; both trends have occurred as the composition of output has shifted away from resource-intensive commodities and as savings in resource use have been achieved in many industries. Moreover, with the exception of forest products, labor-capital costs

per unit of extractive (agricultural, forest, mineral) output have fallen as fast as the cost of nonextractive output since 1870. Then, too, except for forest products, the prices of resource commodities have shown little long-run tendency to rise relative to the prices of all other goods. Hence, the accumulated evidence to date does not substantiate the resource scarcity hypothesis. Apparently, technological change has provided a powerful offset to diminishing returns. In fact, many empirical studies specifying the sources of economic progress have attributed a major proportion of the growth of output per capita to technological change, contrary to the predictions of conservationists and classical economists.

Empirical research on the behavior of fertility patterns also has contributed to economists' growing skepticism of the classical-conservationist doctrine. Studies by Gary Becker, Richard Easterlin, Donald Bogue, and others have shown that (1) parents generally weigh the costs and benefits of having children and make rational decisions regarding births, and (2) that human fertility responds voluntarily to changing environmental conditions. This evidence suggests that the Malthusian assumption of a population growing exogenously at a constant biological rate was wrong. Instead, the population growth rate is now seen as an endogenous variable influenced by economic and social forces. If fertility adjusts to environmental pressures, then, as Mill had pointed out in the mid-nineteenth century, population growth could slow or halt before resource ceilings were reached.

Economic Analysis of Conservationist Claims

Paralleling developments on the empirical front in the 1950's and 1960's were studies by economists Harold Barnett and Chandler Morse, S. Ciriacy-Wantrup, H. Scott Gordon, and Anthony Scott. These studies critically appraised conservationist arguments and demonstrated how economic analysis could be applied to problems of natural resource management. The studies also revealed wide discrepancies between economic and conservationist concepts of efficiency in resource allocation and specified several errors in conservationist reasoning.

First, it is not necessarily true, as claimed by some conservationists, that the private-property, profit-maximizing, free market system promotes a willful disregard for the future and a wasteful spoliation of resources. Economists pointed out that, to the contrary, rational profit-maximizing resource owners with accurate foresight would compare the present value of future gains from resource preservation with the potential current returns from resource exploitation in determining the most profitable use of their resource assets. If society values the future availability of resource stocks more than it values the uses of currently converted resources, then the present value of (or, alternatively, the expected rate of return on) resource preservation would exceed that of resource exploitation. In this case, profit-maximizing owners would be induced to practice conservation. On the other hand, if society values virgin resources less than transformed or converted resource wealth, it would be socially beneficial, as well as privately profitable, to exploit resource stocks. The private property, laissez-faire market system *per se* is not necessarily responsible for excessive and premature resource exploitation. In fact, the opposite may be true. Many economists think that it is when certain vital ingredients necessary to the proper functioning of the market mechanism—knowledge, information, rational behavior, property rights—are absent that non-optimal use of resources may occur. For example, premature and wasteful resource use may stem from (1) irrationality (non-profit maximizing behavior), (2) ignorance of the most profitable opportunities for resource use or uncertainty about future resource prices, (3) imperfections in the capital market that cause resource owners to discount the future at rates higher than those reflecting the true opportunity costs of resource conservation, and (4) externalities or spill-over effects of resource use not taken into account by individual resource users. Government intervention may be justified in such cases.[2] But the proper immediate policy objective, some economists hold, should be correction of the market imperfections rather than attainment of physical resource conservation itself.

Common Property Resources The most important cause of wasteful resource use, many economists believe, has been the incomplete specification of private property rights to scarce natural wealth. Absence of property rights has been characteristic of so-called "common-property" resources such as underground oil pools, oceanic fishing grounds, streams, lakes, and the atmosphere. Since no one possessed property rights to these resources, no one had the incentive to protect them. In such cases, there would be no sole owner to limit access to and prevent excessive use of the resources. Instead, resources would be treated as a free good ("Nobody's property is everybody's property") and thus overused. People would race

[2] It should be noted that some analysts think that certain government policies may *promote* a greater than optimal rate of resource use. Examples of such policies include depletion allowances, tariffs and quotas on raw material imports, unrestricted "free" access to national parks, and price regulation in the natural gas industry where regulatory authorities establish price ceilings at levels too low to permit producers to capture scarcity rents.

to exploit the resource while there was some left. With free access, the resource would appear costless to individual users, all of whom would endeavor to extract the satiation amount. Yet there *would* be real costs of using resources; in this case, the more each individual would take the less there would be available for other users (current and future). Putting resources to a particular use requires the foregoing of alternative uses; other opportunities have to be sacrificed. But since there is no price to indicate the true opportunity cost of resource use, the individual resource user has neither a way of perceiving the cost he imposes on others nor an incentive to conserve the resources for higher-valued future uses. The result: an inefficient pattern of resource exploitation in which higher-valued future uses would be sacrificed for lower-valued present uses. (In fact, in the "free good" case where resource use appears to be costless, some resources may be allocated at the margin to near-zero-valued current uses.) In the jargon of economics, there is an *external diseconomy,* i.e., a divergence between private and social cost.

If the resource could be appropriated as private property, then the owner could charge a fee for its use, thereby restricting access to those willing to pay the fee. Moreover, it would be in the owner's best interest to charge the fee that maximized the economic rent on the resource. This rent-maximizing fee would effectively raise the private cost of resource use, thereby bringing it into equality with the social cost. With resource users now paying full cost, externalities would be internalized, and resource use would be limited to the socially optimal amount. This optimal solution could, of course, be obtained via government regulation; but only if the government acted *as though* it were the sole owner and levied the rent-maximizing user fee. In sum, economists have demonstrated that private property, far from being obstructive of optimal resource management, is conducive to it.

Economists have also pointed out that it is not necessarily true that an endowment of unexploited natural resources is the best legacy the present can leave to the future. Over time, a given resource endowment may be rendered worthless by changes in technology. Even disregarding this possibility, however, it still does not follow that posterity will be made better off by a bequest of virgin natural resources rather than alternative forms of wealth. Those idle natural resources might be more socially productive if converted into other types of wealth, e.g., plant and equipment or investment in research, education, training, and other forms of human capital. If so, then resource preservation would indeed be purchased at a high price, namely, the costly sacrifice of these latter forms of wealth plus the additional wealth they might have produced. To summarize, conservation is just one form of capital investment or provision for the future. As such, it should take priority when it is more productive than other kinds of capital formation. However, if the yield on alternative forms of wealth is higher than on natural resource assets, posterity will benefit more if the present invests in resource conversion rather than in resource preservation.

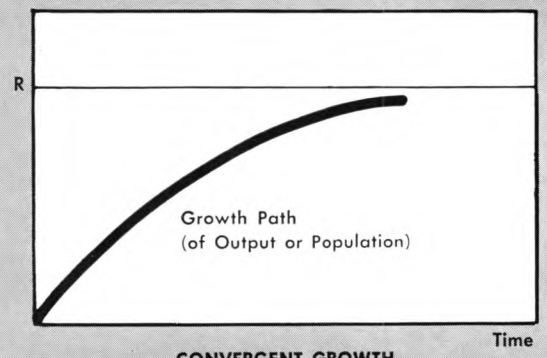
The Reversal of Economic Opinion These critiques of conservationist tenets, like the empirical evidence that contradicted the Ricardian predictions, weakened economists' adherence to the resource scarcity doctrine. Each new study that appeared in the 1950's and early 1960's contributed to a mounting skepticism. By the mid-1960's, many economists had abandoned the discredited doctrine. The switch was now complete. Economists had become skeptics, rather than adherents, of the resource scarcity doctrine.

IV. THE CURRENT DEBATE

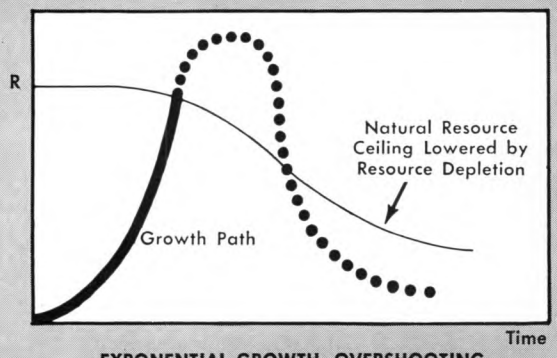
Doomsday Computer Models The resource scarcity doctrine survives today in the computer models that pessimists employ in projecting the future. These models consist of systems of interrelated dynamic equations capable of simulating, from given initial assumptions, future time paths of population, pollution, production, and other economic-demographic variables. Embodied in the structure of the models, albeit in the form of interdependent behavioral relations rather than simple assumptions, are most of the key classical and conservationist postulates—including Malthusian exponential growth of population, Ricardian fixed limits on the supply of land, rapid rates of resource depletion, and irreparable ecological destruction in the form of exhaustion of the limited pollution-absorbing capacity of the environment. Moreover, the computer models follow conservationist tradition in (1) stressing the physical availability (rather than the economic quality) of natural resources, (2) de-emphasizing the impact of the continuous exponential trend of scarcity-offsetting technological progress, and (3) neglecting or overlooking the resource-allocating function of the market price system. Thus, in the computer simulations, time does not alter the types of material inputs required in the production process, technology does not stay the arrival of the ecological Day of Judgment, and no price system exists to in-

Chart 3

ALTERNATIVE GROWTH SYSTEMS



CONVERGENT GROWTH



EXPONENTIAL GROWTH, OVERSHOOTING, AND SUBSEQUENT COLLAPSE

The left-hand diagram depicts a case where growth-retarding feedback mechanisms operate to produce a stable, asymptotically convergent growth pattern. No such stable behavior is evident in the system illustrated in the right-hand diagram, however. Here growth-inhibiting feedbacks are too weak or operate with too long a lag to prevent the accelerating growth path from overshooting its sustainable limits. Note also the falling resource ceiling corresponding to the assumption that the resource-consuming effects of growth exceed the resource-creating effects of technological change. Growth is followed by collapse as the system seeks its greatly reduced sustainable limits.

duce substitution, invention, and other resource-economizing and resource-creating responses to particular shortages.

Alternative Growth Systems The doomsday computer models do differ from the classical Ricardian system in one important respect, however. In the Ricardian model, growth terminates in the stationary state, with population, capital, and output stabilizing at their limit magnitudes. By contrast, in the doomsday computer simulations there is no final stationary equilibrium; instead, growth ends in collapse, with population and output plunging uncontrollably downward.

This difference in the dynamical behavior of the two models stems from the difference in their structures. The Ricardian model contains a stabilizing feedback mechanism that is absent from the computer models. Specifically, in the Ricardian model, growth produces sharply diminishing returns that act as a brake on further growth. By forcing down wage rates and capital yields, such diminishing returns induce reductions in the rate of expansion of labor and capital supplies. Moreover, the closer the system is to equilibrium, the stronger is the growth-retarding force of diminishing returns. Consequently, growth slows to zero as resource limits are approached, and the system converges smoothly on its stationary position.

No such stabilizing adjustment mechanisms operate in the doomsday computer models; or rather they operate too weakly and with too long a delay to prevent disaster. Thus, in the computer simulations, growth tends to be explosive and accelerating rather than dampened and decelerating. Instead of asymptotic, convergent sequences, the computer generates exponential growth paths that collide with or overshoot the resource barriers and then rebound downward rapidly. In sum, a constrained growth system can end in either a stable, sustainable equilibrium or in reversal and decline, as shown in Chart 3. It all depends on the interrelationships among the variables, together with their time-lag structures. If the growth-slowing interrelationships are strong and operate with but slight delay, the growth system will be convergent and self-regulating. But if the growth-inhibiting feedbacks are weak and/or operate with long lags, then the growth system will tend to overshoot its sustainable limits. And if these limits themselves are shrinking because of the resource-consuming and resource-exhausting effects of economic expansion, then growth must give way abruptly to sharp decline. This latter unstable pattern is characteristic of the time paths generated by the current doomsday models. That is why the prospect these models offer, i.e., catastrophic collapse, is more terrifying than the Ricardian prognosis of dull stagnation.

The Stationary State as a Policy Objective Although a stationary state does not exist in the pessimists' computer models of the unregulated world

economy, it is prominent in their policy proposals. But the stationary state advocated is of the Millian, rather than the Ricardian, variety. Since the computer simulations indicate that natural resource constraints do not produce a zero-growth state, then one must be imposed by human decision, i.e., deliberate policy action. In fact, the pessimists present this as the *only* policy solution. In order to prevent catastrophe, they argue, the zero-growth state must be imposed immediately. All other measures are useless, serving at best to postpone the collapse for a few years. For example, the computer indicates that even if population growth were halted, land and raw material productivity quadrupled, usable resources doubled, and pollution cut by three-fourths, apocalypse would nevertheless occur sometime in the next century because of the continued exponential growth of output. Thus, the only recipe for survival is to stop *all* growth—output as well as population.

Halting output growth need not result in an arresting of human progress or betterment, argue the zero-growth advocates. Like J. S. Mill, the no-growth proponents view the self-imposed stationary state as one in which much human activity could be diverted from industrial production to intellectual and cultural pursuits. Moreover, similar to Mill, they recognize that resource-saving, pollution-abating technological progress could raise the quality of life in the stationary state.

Economists' Critique Economists are skeptical of the pessimists' computer models, which they believe to be inaccurate representations of real-world dynamic structures. They argue, for example, that technological progress is inadequately handled in the doomsday models, just as it was in the earlier Ricardian and conservationist schemata. An adequate conception of technological progress, they claim, should recognize its unique scarcity-alleviating characteristics. First, technical knowledge has tended to expand continuously at an exponential rate without suffering diminishing returns. Although the rate of technical change may eventually diminish as it becomes more and more difficult to obtain additional knowledge, many economists expect this to happen only in the far distant future. In the foreseeable future, however, there may be no retardation in this source of growth. Such technological progress could continue to stimulate output growth at a constant rate even though labor, capital, and natural resources had reached zero-growth positions.[3]

[3] Economists have constructed plausible theoretical models in which conventional inputs—labor, capital, natural resources—can be held constant, yet per capita output will expand at the same percentage rate as so-called "factor-augmenting" technological change.

Second, technological progress may be *resource-generating,* i.e., it may create new uses for formerly worthless substances, make available heretofore inaccessible stocks, and make feasible the extraction of formerly high-cost, low-quality materials. In the view of many analysts, the supply of natural resources is not permanently fixed but rather is an increasing function of the level of technical knowledge. The resource-augmenting powers of technical progress, economists argue, may act to raise resource ceilings.

Third, technical advance improves the potential for economizing in the use of particular resources by expanding progressively the possibilities for resource substitution. Doomsday computer models, economists claim, fail to capture these continuous, exponentially-expanding, quality-altering, resource-creating aspects of technology. Instead, computer analysts typically introduce technological change into their models in the form of one-shot adjustments to the initial resource inventory (e.g., an assumed one-time doubling of resource stocks) or to output yields per unit of material input.

But the major deficiency of the computer models, critics hold, is their failure to allow for changes in human behavior. Mechanisms that signal the need for change and then induce the appropriate reactions are noticeably absent in the computer models. Instead, behavior patterns are represented by unchanging parameters whose magnitudes are estimated empirically from past data. Mechanical extrapolation of equations containing these parameters then yields the exponential growth paths of the doomsday simulations.

Adjustment Mechanisms Economists, however, are convinced that things do not work that way. Behavior patterns do not remain invariant in the real world; instead they change, usually in a stabilizing manner. Adjustment mechanisms built into the economy insure this. When economic conditions change, feedback mechanisms provide the signals and incentives that induce stabilizing responses. People react to the signals by adjusting their behavior patterns. For example, when a particular resource becomes scarce, the consequent rise in its price signals shortages and activates the appropriate responses. Users endeavor to economize on the high-priced resource, perhaps by using substitutes or by developing new resource-saving techniques. Resource producers, too, find the rising prices provide a strong incentive to expand supply, via recycling or extraction from lower-grade sources. Moreover, speculators who held the resource idle in anticipation of the higher

price now make it available. All these responses tend to reduce the resource shortage.

Stabilizing adjustments also would tend to occur if *all* resources were becoming scarce, as in the Ricardian model. With growing total resource scarcity, diminishing returns would lower the rate of return on capital, thereby inducing capital suppliers to reduce the rate of natural resource-absorbing capital accumulation.

Adjustment mechanisms also have obviously been at work in altering U. S. fertility patterns, although admittedly economists have been unsuccessful in specifying precisely what those mechanisms are. Fertility has fallen by more than half since 1800, apparently in response to increases in affluence, education, urbanization, and availability and knowledge of birth control techniques. Although the long-run downward trend was interrupted by an upsurge in the 1940's and 1950's, the descent of the birth rate was resumed in the late 1950's. Since then it has continued to decline sharply. The falling birth rate —which is now very close to the zero population growth level—has made a mockery of naive Malthusian extrapolations of constant biological growth rates.

These are just a few of the stabilizing signal-and-response mechanisms that, economists believe, would automatically retard and halt growth long before resource limits were reached. In sum, economists contend that these adaptive mechanisms would transform explosive growth paths into dampened, asymptotically convergent ones.

Correcting Defects in the Price System On the issue of ecological deterioration, economists concede that pollution and environmental problems have been getting worse. But they think it in error to blame the price system for these problems, as some of the more extreme anti-growth spokesmen have done. Economists point out that pollution problems arise not because the price system is evil *per se* but because its coverage does not extend to nonappropriable "common property" resources, e.g., streams, lakes, oceans, the air, etc. In the absence of private property rights, these resources, although scarce, are treated as free goods and thus overused. Thus, pollution proceeds without a cost constraint when there is no price-cost adjustment mechanism to signal the rising scarcity of environmental resources. However, economists believe that environmental degradation can be controlled by the use of pollution taxes or fees (so-called "licenses to pollute"). Such surrogate prices levied on the users of common property resources would transform social costs into private costs (by "internalizing the externalities") and induce the optimal use of those resources.

Costs of a Zero-Growth Solution In general, economists oppose the pessimists' zero-growth prescription for minimizing the harmful effects of progress. Halting growth, they note, would not necessarily stop environmental pollution, which is a function of congestion and the geographical concentration of population rather than of growth. Furthermore, such a drastic solution would be far too costly, they believe. All the potential benefits from growth necessarily would be sacrificed. These sacrifices could indeed be formidable. For example, zero world growth would mean that underdeveloped countries would be frozen at their present levels of poverty. And in the developed countries, zero growth could virtually eliminate the social and economic mobility that has been the main avenue of progress and achievement for many. Geared to change, the whole institutional structure of countries like the U. S. would suffer severe strain if zero-growth policies were imposed.

Moreover, the extreme zero-growth solution is unnecessary, economists believe. Less severe remedies for regulating growth are available. These remedies would permit society to reap the benefits of growth while minimizing harmful side effects. Specifically, defects in the pricing system could be corrected so that the true social costs of growth would appear in private cost calculations. This internalization of all external costs would force private individuals to bear the full costs of their actions. It would also provide strong incentives for cost-minimizing choices. On both grounds, private decisions under the corrected price system would be expected to produce a more nearly socially optimal pattern of growth.

Finally, although economists agree with biologists that growth cannot continue indefinitely, some of them nevertheless think it may persist for a long time —perhaps several centuries—before ceasing. *Particular* resources may run short from time to time, to be sure. But with technology continuing to create new possibilties for substitution, economists see little long-run danger of *general* resource shortage. As in the past, the problem is less likely to be one of complete resource exhaustion than one of necessary resort to progressively higher-cost, lower-grade resource supplies that still exist in abundance. Many economists think that when growth ultimately does cease, it will be in the form of a gradual slowing down rather than a violent collapse. Moreover, in the final no-growth state, the levels of population, per capita incomes and wealth, and technology may far exceed the eve-of-destruction levels predicted by the computers.

Thomas M. Humphrey

REFERENCES

I. General

Barnett, Harold J. and Chandler Morse. *Scarcity and Growth: The Economics of Natural Resource Availability.* Baltimore: Johns Hopkins Press, 1963. The best and most complete examination of the historical natural resource scarcity doctrine. Part I surveys the classical doctrine of Malthus, Ricardo, and Mill and reviews the writings of the Conservationist Movement (1870-1920). Part II is devoted to the construction of analytical growth models embodying the classical doctrine of diminishing returns. Part III contains extensive empirical tests of the hypothesis that natural resource scarcity has been increasing in the U.S. since 1870. The hypothesis is rejected.

Gordon, H. Scott. "Today's Apocalypses and Yesterday's." Unpublished paper delivered at the 85th annual meeting of the American Economic Association, Toronto, Canada, December 27-30, 1972. A brilliant, fascinating sketch of different concepts of economic apocalypse. Compares the current doomsday scenarios with the gloomy speculations of Marx, Jevons, the classical economists, and early twentieth century eugenicists. The distinction between exponential and convergent growth is developed in this paper, as are the stability and adaptability implications of alternative constrained growth systems.

Spengler, Joseph J. "The Economist and the Population Question," *American Economic Review*, 46 (March 1966), 1-24. Surveys economists' views on population, diminishing returns, and resource scarcity from Malthus to the present. Discusses the factors responsible for the post-1930 reversal of economic opinion of the danger of Malthusian overpopulation.

II. Classical Economists and the Natural Resource Scarcity Doctrine

Good discussions of the Malthusian population mechanism, the law of diminishing returns, the Ricardian growth model, and the stationary state (both the Ricardian and Millian versions) may be found in the following sources.

Adleman, Irma. *Theories of Economic Growth and Development.* Stanford, California: Stanford University Press, 1961. See Chapter 4, which deals with Ricardo.

Baumol, William J. *Economic Dynamics.* 3rd ed. New York: MacMillan, 1970. See Chapter 3 for a discussion of the classical dynamics.

Blaug, Mark. *Economic Theory in Retrospect.* Revised ed. Homewood, Illinois: Richard D. Irwin, Inc., 1968, pp. 68-96, 214.

McKinley, Erskine. "The Theory of Economic Growth in the English Classical School." *Theories of Economic Growth.* Ed. Bert F. Hoselitz. Glencoe, Illinois: Free Press, 1961. See Chapter 3.

Robbins, Lord Lionel. *The Theory of Economic Development in the History of Economic Thought.* New York: St. Martin's Press, 1968. pp. 10-14, 28-43, 165-70.

III. Conservationism and the Problem of Natural Resource Depletion

Each of the sources listed below provides an economic analysis of conservation issues. Most of the sources also discuss problems of optimal pricing and optimal use of "common property" resources.

Alchian, Armen A. and William R. Allen. *University Economics.* 2nd ed. Belmont, California: Wadsworth Publishing Co., Inc., 1967, pp. 451-58.

Gordon, H. Scott. "Economics and the Conservation Question." *Journal of Law and Economics*, 1 (1958), 110-21.

Milliman, J. W. "Can People Be Trusted With Natural Resources?" *Land Economics*, 38 (August 1962), 199-218.

Scott, Anthony. *Natural Resources: The Economics of Conservation.* Toronto: University of Toronto Press, 1955.

———. "Conservation: Economic Aspects." *International Encyclopedia of the Social Sciences.* Vol. 8. Ed. David L. Sills. New York: MacMillan and Free Press, 1968, 279-90.

IV. The Current Debate

Two texts that present the doomsday computer scenarios are:

Forrester, Jay W. *World Dynamics.* Cambridge, Massachusetts: Wright-Allen Press, 1971.

Meadows, D. H., D. L. Meadows, J. Randers, and W. W. Behrens, III. *The Limits to Growth.* New York: Universe Books (A Potomac Associates Book), 1972.

Critical reviews by economists of the doomsday scenarios are:

Anderson, David C. "A Careful Look at Growth as Suicide." *Wall Street Journal*, March 17, 1972.

Barnett, Harold J. Review of Forrester's *World Dynamics. Journal of Economic Literature*, 10 (September 1972), 851-54.

Kaysen, Carl. "The Computer That Printed Out W*O*L*F*." *Foreign Affairs*, 50 (July 1972), 660-68. A superb review article containing especially good explanations of adjustment mechanisms, exponential vs. convergent growth paths, the nature of technological change, and physical vs. economic concepts of resources.

"Limits to Misconception." *The Economist*, March 11, 1972, pp. 20-21.

Shubik, Martin. "Modeling on a Grand Scale." *Science*, 174 (3 December 1971), 1014-15.

Somewhat more sympathetic reviews are:

Benoit, Emil. "Must Growth Stop?" *Columbia Journal of World Business*, 7 (May-June 1972), 41-44.

Heilbroner, Robert L. "Growth and Survival." *Foreign Affairs*, 50 (October 1972), 139-53.

Some recent writings expressing skepticism about claims of impending resource scarcity and imminent environmental disaster are:

Barnett, Harold J. "Population Problems—Myths and Realities." *Economic Development and Cultural Change*, 19 (July 1971), 545-59.

Easterlin, Richard A. "Does Human Fertility Adjust to the Environment?" *American Economic Review*, 61 (May 1971), 399-407.

Nordhaus, William and James Tobin. "Is Growth Obsolete?" *Economic Growth.* Vol. 5 of Fiftieth Anniversary Colloquium Series. New York: National Bureau of Economic Research, 1972, 1-80.

Schultz, Theodore W. "The Ecosystem Doom.' *Bulletin of the Atomic Scientists*, 28 (April 1972), 12-17.

Wallich, Henry C. "How to Live With Economic Growth." *Fortune* (October 1972), pp. 114-16, 121-22.

BANK AFFILIATES AND THEIR REGULATION: PART I

Until 1933, there were no Federal laws specifically designed to limit and regulate relationships among banks and other business organizations under common ownership or control. Today, as a consequence of the Banking Act of 1933 and the Bank Holding Company Act of 1956, banks and other companies related by common ownership or control are known as "affiliates," and a statutory framework has come into existence that governs not only the extent of permissible affiliation but also many important aspects of the affiliate relationship. The background, rationale, and current status of Federal regulation of banks and their affiliates are reviewed in this article, with particular emphasis upon affiliation by means of bank holding company ownership.

INTRODUCTION

From very early in the history of the United States, commercial banks have been confined by statutes, court decisions, public opinion, and their charters to the performance of a relatively few specified activities at certain authorized locations. Permissible activities consist primarily of receiving deposits, making loans and investments, furnishing trust services, and performing a number of functions incidental to authorized activities. The locations of bank offices are controlled by Federal and state laws regulating the establishment of branches.

The activities of commercial banks are restricted because they operate to a very large extent with funds owned by the public. When people deposit money with a bank, they do not expect it to be endangered because the institution is engaged in speculative ventures. People also expect to be able to withdraw their funds on demand or on relatively short notice. Not only the liquidity but even the solvency of banks could be affected if they were permitted to engage in general business and financial enterprises or to tie up their funds in manufacturing, distributing, or retailing operations.

The policy limiting the establishment of branches stems from other considerations. One of the most important of these involves the preservation of the dual banking system in the United States, a system based upon the existence of national banks and state-chartered banks competing, side-by-side, within each state throughout the country. Maintaining the uneasy balance of power between these two banking systems has required that branching by national banks be permitted only within the boundaries of the particular state where the head office of the national bank is located, and then only to the geographical extent that state-chartered banks may establish branches within the state.

But even though no restrictions were imposed prior to 1933 on the extent to which individuals, corporations, and other organizations owning banks could diversify their operations by acquiring other banks or by engaging in nonbanking activities, experiences with uncontrolled affiliation during the speculative boom of the 1920's led to public demand for the regulation of relationships among banks and their affiliates. During this period, bank funds belonging to depositors were used to finance the operations of affiliated corporations, often with large losses. The problem was summarized in these words by a Senate Subcommittee in 1931:

> **Basically, there can be no objection to the stockholders of a bank engaging in any other business they prefer with their own funds. However, if such activities tend to affect directly the position and soundness of the bank itself, they then become of prime importance in the regulation of banking.[1]**

Less than two years later, the Banking Act of 1933 entirely separated common ownership and control of member banks of the Federal Reserve System from securities companies. Other statutory provisions enacted in that year imposed strict limitations on financial dealings between member banks and their affiliates and subjected affiliates to examination and reporting requirements of Federal bank supervisory authorities.

The first statutes designed to regulate bank holding companies were also enacted in 1933. However, except for securities companies, the law did not require separation of control over banks and nonbank businesses, nor did it require prior approval by any Federal authority for acquisitions of voting shares of banks. The only control over bank holding companies was through the voting permit mechanism. In order to vote shares of member banks, bank holding companies were required to obtain voting permits; by this means some measure of control was exercised

[1] "Operation of the National and Federal Reserve Banking Systems," Hearings pursuant to S. Res. 71, 71st Cong., 3rd Sess. (1931), p. 1063.

even over bank affiliates that were not members of the Federal Reserve System.

Major changes in the regulation of bank holding companies were made with enactment of the Bank Holding Company Act of 1956 and its amendments in 1966 and 1970. These statutes completely replaced the 1933 provisions applicable to bank holding company affiliates. Today, the extent of permissible affiliation with banks by means of bank holding company ownership or control is closely regulated, although affiliation through ownership or control other than by bank holding companies remains subject only to the 1933 restrictions that prohibit affiliation of member banks with securities companies. In addition, all banks insured by the Federal Deposit Insurance Corporation must now comply with certain affiliate provisions, regardless of their membership in the Federal Reserve System or their ownership by bank holding companies.[2]

Present law relating to banks and their affiliates is thus a product of two different but not entirely unrelated events. The first of these was the attempt by a number of banks, commencing about 1908, to expand into the investment banking business by organizing affiliate corporations. The second was the growth of bank holding companies. The result of these events was enactment of two different sets of statutory provisions governing affiliate relationships. The first of these, which came into being in 1933, applies to all types of affiliates. The second set of laws, enacted in 1956 and amended significantly in 1966 and 1970, is applicable to bank holding companies and their subsidiaries.

This complex and not altogether logical regulatory structure can best be understood by tracing its growth and development from the inception of the affiliate movement.

ORIGIN OF BANK AFFILIATE REGULATION

The formation of bank affiliates began in 1908 when the president of the First National Bank of New York sent a letter to the bank's stockholders explaining why an affiliated corporation was to be organized. The letter stated, in part:

> Dear Sir:
>
> It is deemed to be for the best interests of the stockholders of this bank that a security company, such as has proved advantageous in the case of several other banks, should be organized for the purpose in part of transacting for its patrons certain lines of profitable business, which, though often transacted by bankers, are not expressly included within the corporate powers of national banks. Among these are the acquiring and holding of real estate, securities, stocks and other property.[3]

Shortly afterward, the National City Bank of New York also established an affiliate. By 1920 a number of other national and state-chartered banks had affiliates, organized like those of First National and National City under the general business corporation laws of the states to engage in nonbanking activities that were prohibited to banks themselves under the banking laws.

In the beginning, some banks used their affiliates as vehicles to acquire stock interests in other commercial banks. For example, soon after its organization the affiliate of First National Bank acquired over 50 percent of the stock of Chase National Bank. It also held substantial stock in the First National Bank of Minneapolis, the Minneapolis Trust Company, the Astor Trust Company, the Bankers Trust Company, the Brooklyn Trust Company, the Liberty National Bank, the National Bank of Commerce, and the New York Trust Company. As a consequence of similar acquisitions, a controversy developed between the Department of Justice and National City Bank's affiliate, the National City Company, over its ownership of stock in 16 banks, reportedly the largest concentrated holding of bank stock in the United States at that time.[4] However, this activity came to an end in 1911 when the Solicitor General of the United States issued an opinion (concurred in by the Attorney General) holding that national banks could not legally acquire and hold the stock of other national banks by means of affiliated corporations. Thereafter, the question of unregulated acquisitions of bank stocks by banks and bank holding companies did not recur until late in the 1920's.

The Growth of Securities Affiliates In contrast, nothing was done prior to 1933 to impede the use of bank affiliates to engage in nonbanking activities or to acquire nonbanking companies. Particularly rapid growth occurred among securities affiliates in response to the enormous financing requirements of the United States and allied governments during World War I. In August 1916, the National City Company took over the bond department of its affiliated bank, at the same time purchasing the entire business of a large investment banking company with over 200 employees and offices in Philadelphia, Boston, Balti-

[2] As more fully discussed later in this article, since July 1, 1966, the provisions of Section 23A of the Federal Reserve Act (80 Stat. 242), dealing with loans and other financial relationships between a member bank and its affiliates, apply to every nonmember insured bank just as if it were a member bank. Similarly, affiliates of nonmember insured banks are subject to examination by the Federal Deposit Insurance Corporation (80 Stat. 1053).

[3] "Operation of the National and Federal Reserve Banking Systems," Hearings pursuant to S. Res. 71, 71st Cong., 3rd Sess. (1931), p. 1053.

[4] Peach, *The Security Affiliates of National Banks* (1941), p. 144.

more, Washington, D. C., Cleveland, Detroit, Buffalo, Albany, San Francisco, and Pittsburgh. By 1929, there were at least 132 securities affiliates of national and state-chartered banks, and at least 459 banks were themselves engaged in the securities business.[5] According to information presented to a Senate Subcommittee in 1931, commercial banks and their affiliates accounted for almost 45 percent of the total originations of bonds in 1930, compared with only about 22 percent in 1927.[6] The same group of banks and securities affiliates participated in 61.2 percent, by dollar value, of the total volume of bond distributions in 1930, compared with only 36.8 percent three years earlier.[7] Commercial banks and their affiliates were thus in process of taking over the bulk of the investment banking business in the United States when the stock market collapsed in October 1929.

Apart from their position as underwriters and distributors, securities affiliates engaged in numerous other activities in the 1920's, among them the following:

> (1) Retailing securities, including maintaining corps of salesmen and branches in states other than that in which the affiliated bank operated;
> (2) Acting as holding companies to carry blocks of securities for control that the affiliated banks could not (or preferred not to) list among their investments;
> (3) Acting as investment trusts, buying and selling securities acquired purely for investment or speculative purposes;
> (4) Functioning as assets realization companies to take over from affiliated banks doubtful or nonliquid assets, financing these acquisitions in some cases by obtaining loans from the affiliated banks from whom the assets were purchased;
> (5) Providing a medium for supporting the stock of affiliated banks; and
> (6) Operating as a real-estate holding company.[8]

The report of a Senate Subcommittee that investigated securities affiliates in 1931 pointed out that most of the above activities could involve risks of substantial loss to affiliated banks. According to the report, affiliates acting as security holding companies or as investment trusts showed a much greater tendency to operate with borrowed funds than similar organizations that were independent of banks. The reason given for this conclusion was that ". . . the identity of control and management which prevails between the bank and its affiliate tends to encourage reliance upon the lending facilities of the former."[9] Similarly, the Subcommittee found that activities of affiliates in receiving doubtful assets, in supporting the affiliated bank's stock, and in holding real estate caused substantial losses in certain instances.[10] This conclusion was supported by the testimony of a number of bankers who conceded in the course of testifying before the Subcommittee that serious abuses developed in the 1920's as a consequence of financial relationships between banks and their securities affiliates.[11] The Chairman of the Board of National City Bank gave the following testimony:

> Senator Norbeck: You said that affiliates were abused by the parent companies?
> Mr. Mitchell: No. I meant that the development of affiliates as a principle was something that had been abused. We saw it in the case of the Bank of the United States where there was the custom of forming an affiliate to take over any bad loan. I think that the number of their affiliates ran to 50 or 60. They were incorporated in New York and were in truth affiliates. Those corporations took over faulty assets and immediately borrowed from the bank—a very dangerous practice.[12]

The failure of the Bank of the United States in New York City in December 1930 focused national attention upon the possibilities of abuse inherent in the relationship between banks and their affiliates, especially securities affiliates. Although bank failures had been at a high rate throughout the 1920's, the failed banks were typically small, undercapitalized institutions in rural areas of the South and Midwest. These banks were regarded as organizations whose time had passed with the rise of the automobile and the shift of population toward fewer, larger urban areas—banks whose survival was already in jeopardy as a consequence of the agricultural depression that began in 1920 and continued throughout the decade. While a matter of concern to Congress and the supervisory authorities, the bank failures of the 1920's were viewed as symptoms of adjustment to fundamental economic and technological change.

Entirely different questions were raised by the failure of the Bank of the United States. Here was an institution with deposits in excess of $200 million, more than the combined deposits of the 551 banks

---

[5] *Ibid.*, p. 83. Although other sources indicate somewhat different totals for the number of banks engaged in the securities business, it appears likely that the tabulations by Mr. Peach are the most reliable.

[6] "Operation of the National and Federal Reserve Banking Systems," Hearings pursuant to S. Res. 71, 71st Cong., 3rd Sess. (1931), p. 299.

[7] *Id.* The total dollar volume of originations by all companies, including both commercial banks and their securities affiliates and private investment bankers, was $5.9 billion in 1927 and $4.6 billion in 1930. The total dollar volume of bond distributions by the same groups was $13.2 billion in 1927 and $12.8 billion in 1930. The "originating" investment banker is the one that first discusses the particular bond issue with the corporation or governmental unit in need of funds. The "participants" are other investment bankers that agree to share in the risks and benefits of underwriting and distributing an issue of securities. Shultz and Squier, *The Securities Market* (1963), p. 67, fn. 1.

[8] "Operation of the National and Federal Reserve Banking Systems," Hearings pursuant to S. Res. 71, 71st Cong., 3rd Sess. (1931), p. 1057.

[9] *Ibid.*, p. 1058.

[10] *Id.*

[11] *Ibid.*, pp. 307, 342, 404.

[12] *Ibid.*, pp. 306-7.

that had failed in the entire United States during the year ending June 30, 1929, more than the aggregate deposits of all 484 banks that failed throughout the country during the previous 12-month period ending June 30, 1928. Together, the failure of the Bank of the United States in 1930 and Bankers Trust Co. in Philadelphia at about the same time involved deposits in excess of $650 million, an amount greater than the deposits of all 1,035 banks that failed in the nation in the two years ending June 30, 1929.

No doubt some of the bank failures that occurred in the wake of the stock market panic in October 1929 were traceable to abuses involving securities affiliates of banks. Yet failures due to this cause could not have accounted for more than a handful of total bank failures. There were a total of 9,096 bank failures in the years 1930-1934, involving depositor losses of $1.3 billion,[13] but relatively few banks had securities affiliates—probably less than 200.[14]

In reality, most banks that had securities affiliates survived the holocaust of bank failures from 1930 through 1933, although the affiliates sustained substantial losses in a number of cases. After an extensive investigation by the Senate Subcommittee, only the Bank of the United States was singled out as an example of failure caused by the relationship of the bank and its securities affiliate. A report by the Subcommittee had the following comments summarizing operations of 14 bank securities affiliates in 1930:

Results of security affiliate operations:

The financial results of the operation of security affiliates during the period following the stock market collapse were on the whole unfavorable. Losses of any substantial size were not reported in every case, the chief exceptions being those organizations which restricted themselves to the distribution of high-grade bonds.[15]

The Subcommittee did, however, list a number of different ways in which the operations of a securities affiliate could affect the position of the affiliated bank, stating that "(i)n actual practice, the operations of a number of securities affiliates have affected the parent institutions to a greater or lesser degree in several of the ways outlined."[16] These were summarized as follows:

(1) "Very prevalent" borrowing by the affiliate from the bank;

(2) The selling of securities by an affiliate to its bank or other affiliates under repurchase agreements, or their purchase by the affiliate from the bank in the same manner;

(3) The purchase of securities by the bank to relieve the affiliate of excess holdings;

(4) More liberal lending by the bank to customers on issues sponsored by the securities affiliate in order to support their distribution. The Subcommittee report stated ". . . it may prove more difficult to insist upon the maintenance of adequate margins on these security loans than on other such advances, in view of the fact that customers are encouraged to make the loans by the bank's own affiliate;"[17]

(5) Injury to the good will of the bank if depositors suffered substantial losses on securities purchased from the bank;

(6) Causing of undesirably wide fluctuations in the price of the affiliated bank's stock as a result of purchases and sales of the stock by the securities affiliate. The report noted that ". . . efforts made in some cases to push the sale of the bank's stock through the affiliate to depositors of the institution hurts the position of the bank when its shares suffer a major market decline subsequently;"[18]

(7) Making of unwise commitments by the bank, in the knowledge that in case of need they could be shifted to affiliates and thus removed from the bank's condition statement;

(8) In reliance upon the resources of the parent bank in case of need, the tendency of securities affiliates to assume commitments less cautiously than private investment bankers; and

(9) In the case of banks with both a trust department and a securities affiliate, adverse effects upon the independence with which fiduciary activities were exercised.

Relations between banks and their securities affiliates were thus a source of major Congressional concern by 1931. Meanwhile, especially in the upper Midwest and on the West Coast, affiliated bank

---

[13] BANK FAILURES AND DEPOSITOR LOSSES, 1921-1970

	Total Bank Failures	Depositor Losses (in thousands)
1921-1929	5,711	$ 564,731
1930-1934	9,096	1,336,533
1935-1940	445	9,166
1941-1950	52	139
1951-1960	40	1,926
1961-1970	44	33,497

Source: Federal Deposit Insurance Corporation.

[14] According to Peach, *The Security Affiliates of National Banks* (1941), p. 83, the number of national and state-chartered banks operating securities affiliates were as follows in the years shown below:

BANKS WITH SECURITY AFFILIATES

	National	State	Total
1927	60	22	82
1928	69	32	101
1929	84	48	132
1930	107	75	182
1931	114	58	172
1932	104	53	157

As the table indicates, even before legislation was passed in 1933 requiring the divorcement of securities affiliates, many banks had begun dismantling them. The high tide of securities affiliates was in 1930, with 182; but by 1932 the number was down to 157. This decline does not appear to have been due to failures of affiliated banks. Peach states at p. 158, with reference to May 1933, that "(m)any other affiliates were already in process of liquidation, or had been previously dissolved, either because final passage of (The Banking Act of 1933) was anticipated or because banks welcomed the opportunity to rid themselves of affiliates which they had thought necessary or highly desirable during the twenties."

[15] "Operation of National and Federal Reserve Banking Systems," Hearings pursuant to S. Res. 71, 71st Cong., 3rd Sess. (1931), p. 1061.

[16] *Ibid.*, p. 1064.

[17] *Id.*

[18] *Ibid.*, pp. 1063-64. The existence of these and other abuses by certain banks and their securities affiliates was also shown in the course of extensive hearings before the Senate Banking and Currency Committee in 1933 and 1934. Hearings on S. Res. 84, 72nd Cong., 2nd Sess. (1933); Hearings on S. Res. 56 and S. Res. 97, 73rd Cong., 1st Sess. (1933-1934).

holding companies were being used to combine ownership of a variety of nonbanking enterprises with control over extensive regional banking organizations.

Bank Holding Companies and Their Affiliates

Contemporaneously with the growth of securities affiliates, bank holding companies began to form. Although there were several reasons for the growth of bank holding companies between 1927 and 1929, a particularly important factor was the desire of leading bankers in different parts of the country to expand geographically in ways that were prohibited to banks themselves under existing branch banking legislation.

One of the most ambitious plans centered around the announced objective of the Transamerica group in California to establish a nationwide system of branch banks and operate these banks in conjunction with nonbank enterprises. The vehicle to achieve this goal was to be Transamerica Corporation, formed in October 1928, to bring under common ownership several large banks in California and New York, including Bank of Italy, with 294 offices in California; Bank of America of California, with 148 banking offices in the same state; and The Bank of America National Corporation, New York, with 34 banking offices in that state.[19] Other subsidiaries of the holding company included a securities and realty corporation that reportedly owned stock in at least 70 domestic and 59 foreign banks, a security underwriting firm, a mortgage company, a fire insurance company, and two farm loan companies. A report compiled in 1929 by the Research Department of the Los Angeles Stock Exhange included the following glowing appraisal of Transamerica's prospects:

> This great combination is in a position to assure itself profits in many fields. Its banks provide the funds necessary for any desirable deal; its own investment houses underwrite stocks and bonds, which may be marketed and protected by its wide-flung bond houses and securities companies; its banking and investment houses can divert tremendous business to its insurance company; its banking offices may act as agents for the farm loan and real estate mortgage companies; its real estate companies can help to liquidate foreclosed real estate of other departments; its stock trading and brokerage companies have tremendous sources of information and almost unlimited financial support.[20]

Two other leading bank holding companies of this period (ones that were soon to meet with disaster) were the Guardian Detroit Union Group (Inc.) and Detroit Bankers Co. At the end of 1931, these two holding companies controlled more than 55 percent of the aggregate banking resources in Michigan and over 80 percent of all national bank resources in the state.[21] All told, by 1931 there were 97 bank holding company groups, each consisting of three or more banks, operating a total of 978 banks with loans and investments exceeding $8.7 billion.[22] A number of these companies, like Transamerica, were authorized to engage in many different activities and included securities companies among their subsidiaries. It was said, for example, that the securities affiliate of Union Trust Co. of Detroit could do everything under its charter except solemnize marriages and hold religious services.[23]

From many different quarters, concern was expressed over the growth of bank holding companies even before the Depression began in October 1929. Opponents of branch banking saw the holding company as a device to evade Federal and state branching restrictions. Particular alarm was expressed over the announced plan of the Transamerica group to create a national banking organization. Similarly, private investment bankers objected to the ability of a holding company with a network of locally-owned banks to use these banks as outlets to distribute securities to the public, thereby increasing competition and taking away a substantial volume of business from the investment houses.[24]

Apart from these considerations, bank failures occurred among holding company banks in the late 1920's just as among banks generally. On several occasions during these years Congress was advised of risks to holding company banks inherent in the fact that the Federal supervisory authorities had no jurisdiction over holding companies and their nonbank affiliates.

As early as January 1926, the Board of Governors of the Federal Reserve System addressed a letter to Congressman McFadden recommending that there be incorporated in the pending McFadden bill ". . . certain provisions designed to secure adequate information regarding national and state member banks which are closely related in man-

---

[19] "Branch, Chain and Group Banking," Hearings before the Senate Committee on Banking and Currency, 71st Cong., 2nd Sess. (1930), p. 247. Except for Transamerica, other holding company groups apparently were interested only in expansion within certain regions of the country or within a particular trade area. One writer described the bank holding companies of the late 1920's as "separate and isolated growths arising out of the economic life of different and widely separated communities." Collins, *Rural Banking Reform*, pp. 97-8, quoted in Fischer, *Bank Holding Companies* (1961), p. 35. At least some of these local or regional holding companies opposed the Transamerica plan for a national banking organization.

[20] *Supra*, note 19, p. 246.

[21] "Operation of the National and Federal Reserve Banking Systems," Hearings on S. 4115, 72nd Cong., 1st Sess. (1932), p. 121.

[22] "Origin and Development of Group Banking," *Federal Reserve Bulletin* (1938), p. 97.

[23] Hearings on S. Res. 84, 56, and 97, Senate Committee on Banking and Currency, 73rd Cong., 2nd Sess. (1934), p. 4776.

[24] Peach, *The Security Affiliates of National Banks* (1941), p. 103. The large city banks in New York and Chicago were important sources of funds for private investment bankers and did not present the same degree of competitive threat in distributing securities as did the holding company groups with their broad geographic coverage.

agement, operation or interests to other banking institutions and, in particular, to afford some check upon the abuses frequently occurring from chain banking."[25] In its annual report to Congress for the years 1927 and 1928, the Board pointed out that financial companies specializing in the purchase of bank stock were being organized in increasing numbers and that since these companies were not directly engaged in banking, they were not subject to supervision or examination by Federal authorities. Although the Comptroller of the Currency made little comment about bank holding companies prior to 1930, in his annual report of that year he expressed the view that these companies should be brought under the supervisory powers of the Federal Government in those cases where membership in the group included national or state member banks of the Federal Reserve System. He said further that legislation in this respect seemed to be necessary in the public interest.

National Bank Affiliates On March 29, 1932, Governor Eugene Meyer of the Board of Governors of the Federal Reserve System furnished detailed information to the Senate Banking and Currency Committee regarding the extent of affiliation involving national banks. (Many state banks and trust companies had affiliates, but Governor Meyer limited his report to affiliation involving national banks, presumably because of incomplete data regarding state-chartered institutions, the great majority of which did not belong to the Federal Reserve System):

> Senator Glass: From your own experience do you see any necessity for affiliates?
>
> Mr. Meyer: Yes sir, there are a number of affiliates which you in the Congress authorized.
>
> Senator Glass: I mean those unauthorized by law.
>
> Mr. Meyer: Yes.
>
> Senator Glass: We are not talking about these little inconsequential form matters.
>
> Senator Couzens: I should like to ask about affiliates built up outside of the law.
>
> Mr. Meyer: Well, there are 15 different kinds of affiliates.
>
> Senator Couzens: I mean all those not authorized by law.
>
> Mr. Meyer: There are realty companies, holding companies, bank building companies, mortgage companies, liquidating companies, agricultural loan companies, personal or small loan companies, investment trusts, building and loan associations, insurance companies, finance and acceptance cor-

[25] *Supra*, note 19, p. 442.

Table I

NONBANKING AFFILIATES OF NATIONAL BANKS, 1933

Kind of Affiliate	Mode of Control: Stock Owned By Bank	Stock Trusteed	Stock Owned By Other Affiliate	Stock Owned By Bank Stockholders	Total
Securities cos.	4	126	17	45	192
Realty cos.	4	33	8	110	155
Holding cos.	4	28	7	31	70
Bank building cos.	42	3	1	5	51
Safe deposit cos.	33	4	3	4	44
Mortgage cos.	2	11	6	18	37
Liquidating cos.	—	6	3	26	35
Agricultural loan cos.	—	6	3	26	35
Personal loan cos.	1	10	—	16	27
Investment trusts	1	1	7	8	17
Building and loan assns.	—	1	1	14	16
Insurance agencies	—	2	3	10	15
Finance & acceptance corps.	—	2	2	3	7
Title cos.	—	2	1	4	7
Foreign banks	3	—	3	—	6
Joint stock land banks	—	2	1	3	6
Title and mortgage cos.	1	—	2	—	3
Investment houses	—	—	—	1	1
Life or casualty insurance cos.	—	—	—	1	1
Miscellaneous	3	8	13	21	45
Total	98	245	81	346	770

porations, title and mortgage companies, and so on. I have a list here as follows: (see Table I).

Senator Couzens: The most objectionable one on the list is the investment affiliate.

* * *

Senator Couzens: As a matter of fact, practically all of these affiliates were developed by smart attorneys to evade the banking laws.

Mr. Meyer: Yes.

Senator Couzens: Why should they be encouraged to be allowed to continue?

Mr. Meyer: I cannot say, except that they were allowed to continue, and the Congress of the United States knew as well as anybody else that they were developed to evade the law.

Senator Glass: Now, when we want to put a stop to them, we are told that the time is inopportune.

Mr. Meyer: A practice that has grown up with the semblance of legality, even though primarily designed to evade a law, because the authorities and the Congress of the United States did not object to it, although they knew about it, gains a certain right to be treated with some consideration.

* * *

Senator Glass: It has only in recent years come to the attention of Congress that they had been guilty of gross abuses and have created a great deal of distress in the country.

Mr. Meyer: Mind you, Senator, I quite agree, and the board does, with the general purpose and express sympathy with your attitude. In this particular case I think "when" may be important, although I do not know, because we have not the information. I am in hearty sympathy personally, I may say, Senator, with your desire to separate commercial banking from investment banking. That is what you really have in mind.[26]

In his testimony, Governor Meyer emphasized the lack of available information regarding the activities of affiliates other than securities affiliates. For this reason, the Board opposed the complete separation of all affiliates from banks, although it favored the divorcement of banks from securities companies. In a letter to the Committee dated March 29, 1932, Governor Meyer summarized the Board's recommendations as follows:

> With respect to affiliates the Board believes that important reforms to be accomplished at the present time are the granting of power to the supervisory authorities to obtain reports and to make examinations of all affiliates of member banks and the prescribing of limitations on the loans that a member bank may make to its affiliates. The Board realizes that many evils have developed through the operation of affiliates connected with member banks, particularly affiliates dealing in securities. The attached memorandum contains a draft of a provision for the separation of such affiliates after a lapse of 3 years.[27]

William F. Upshaw

---

[26] "Operation of the National and Federal Reserve Banking Systems," Hearings on S. 4115, 72nd Cong., 1st Sess. (1932), pp. 391-94.

[27] *Ibid.*, p. 403.

Next month, Part II will discuss the 1933 affiliate and bank holding company legislation, the 1935 amendments to such legislation, and the events that led to the Bank Holding Company Act of 1956.

The MONTHLY REVIEW *is produced by the Research Department of the Federal Reserve Bank of Richmond. Subscriptions are available to the public without charge. Address inquiries to Bank and Public Relations, Federal Reserve Bank of Richmond, P. O. Box 27622, Richmond, Virginia 23261. Articles may be reproduced if source is given. Please provide the Bank's Research Department with a copy of any publication in which an article is used.*