RESOURCES AND ECONOMIC GROWTH

by Robert M. Solow*

I. Introduction

I shall interpret the topic at hand to mean something quite precise. How much drag on the future economic growth of the developed world is likely to be exerted by the narrowing of its resource base, assuming indeed that the resource base is narrowing? Turned around in time, the same question becomes: what has been the effect on past economic growth of differential availability of natural resources?

This way of posing the problem deliberately excludes certain other angles from which one could approach the relation between resources and economic growth. For example, I shall not discuss the view that the exploitation of an indigenous natural resource, renewable or nonrenewable, is a particularly likely way for the process of modern industrial growth to begin. Directed toward the past, these are the "staple" theories; directed toward the future, they are the various theories of export-led growth. Nor will I discuss questions about the management of their natural resources by currently less-developed countries, nor those about the relation between cheap natural resources and lifestyles in developed countries. Those questions are all interesting, but not the sort that economic theory can very well handle.

Much of the current interest in the relation between natural resources and economic growth had its origin in the splash made by several "doomsday scenarios" a few years ago. Probably for that reason, there is a tendency to pose the question in all-or-none terms: what will happen if the world runs out of essential nonrenewable resources soon? No really defensible answer can be given to that question, as one can see by turning it around in time. To ask what the modern world would be like if there had never been any iron, coal, oil and copper is to ask for science fiction, not science and not history.

In fact, the "running-out" figure of speech is geologically inappropriate in most instances. There is much more copper in the earth's crust than the human race is ever likely to need. Most of it is in ores that are very lean, so lean that they cannot be economically mined and concentrated with currently known technology and the current price of refined copper. In the absence of new discoveries (which are still occurring, even for a metal that has been an object of commerce as long as copper) or new technology, the world economy would move historically along a rising cost-gradient for nonrenewable resources. As the richer ores were used up, the cost of minerals and metals would rise because more labor and capital (and minerals and metals) would have to be expended per unit of refined product. Less valuable uses of each resource would be pared away as the cost rises, or other cheaper materials would be substituted for scarce natural resources, eventually use of a given material would cease, or be limited to the flow that could be made economically available through recycling.

Of course this hypothetical scenario has not worked itself out before our eyes, because new discoveries have been made and, what is much more important, technological improvements in mining, processing and transporting have continued to occur. (At the turn of the century, the leanest copper ore that could be exploited economically was one containing about 5 percent copper. Today 0.5 percent copper-bearing ore can be profitably mined, with essentially no increase in the "real" price of refined copper.)

With this picture in mind, the basic question can be rephrased. What difference would it make to the past (or future) growth of a modern economy if the cost-gradient for important natural resources had been (or were to be) a little steeper or a little higher? In that form, one can hope to give some sort of an answer.

One further preliminary explanation is called for. I shall take an aggregative approach to the question just posed. That is to say, I want to talk about the effects on aggregate real output of somewhat easier or harder access to nonrenewable resources in general. For some purposes one would want to be much more specific, both on the output side and on the input side. How would transportation have developed had coal been cheaper or oil

* Institute Professor of Economics, MIT

more expensive? How would housing or urban life in general have been different if steel for structures and appliances had been more scarce? There may even be some minor but critical resources whose effective disappearance might make a drastic difference to the way we live. But I shall ignore such questions in favor of drab talk about real GNP and real resource use.

II. Analytical Framework

In a commonsense way, one would naturally say that the quantitative relation between resource availability on the one hand and the level and rate of growth of real output on the other depends mainly on three factors:

(a) the "importance" of natural resource inputs;

(b) the degree of substitutability between nonrenewable resources and other inputs in the production of final output;

(c) the pace and bias of technological progress. That is to say, a narrowing of the resource base or a rise in resource costs will have a drastic effect on the level and growth of final output if nonrenewable resources play a very large role in production, if renewable resources and labor cannot easily be substituted for them, and if the course of resource-saving technological progress is slow.* In the opposite case, the increasing cost of nonrenewable resources will have little effect on the level and growth of real output because they do not matter very much in the first place, or because the need for them can easily be circumvented by substitution or invention. The task of theory is to lend more precision to these general remarks.

The first two factors mentioned can most directly be elucidated in a static framework. After that has been done, the results can be incorporated in the standard sort of growth-accounting framework, within which the third factor plays a natural role.

Here is the simplest sort of macroeconomic model which seems to offer any chance of a serious answer to the question at hand. Imagine an economy which produces a single final good called "real output" using, as inputs, capital goods, labor, and a flow of a composite nonrenewable resource, with constant returns to scale. There is a production function that relates any bundle of labor, capital and resource inputs to the largest flow of real output that bundle is technologically capable of producing. Presumably the production function allows some scope for a given output to be produced with more or less consumption of resources, provided less or more labor and/or capital is available. If there were no possibility of substitution, we would not need any analysis. But the facts show very clearly that some substitution is possible.

(Two side comments are in order here. First, one of the ways in which other inputs can be effectively substituted for nonrenewable resources is by a change in the composition of real output to include fewer resource-intensive goods and more of other goods. In a disaggregated model these substitutions in consumption could be treated directly; here they have to be regarded as substitutions in production. Second, common observation suggests that considerably more substitution is possible in the long run than in the short run. For example, fuel-saving opportunities given the existing stock of automobiles are much more limited than in a longer run when the design of automobiles has had time to adapt to new circumstances, and the stock of cars has had time to turn over. This should probably be taken as a long-run model, with the recognition that short-run adjustments might be more difficult.)

Now imagine that the economy has to "buy" its resource inputs at a given price in terms of final output. These purchases could be actual imports, of course, in which case the price represents the going terms of trade between resources and industrial commodities. But the resources could just as well be indigenous; then the price represents the real cost of extraction. In either case the problem to be studied is precisely: What is the consequence for *net* consumable output of a rise in the price of resource inputs? If gross output is called $Q = F(K, L, R)$, where F is the production function and K, L, R are the current flows of capital and labor services and natural resources, then net output $Y = Q - pR$, where p is the price of a unit of resources in the terms of final output. Narrowing of the resource base means higher-cost resources. What then is the effect on Y of a higher value of p?

The answer is surprisingly simple. If the economy reacts so as to maximize its net output, then the elasticity of net output with respect to p is $-s/(1-s)$ where s is the elasticity of gross output with respect to resource input. If the production function

* William Hogan and Alan Manne have independently used an analogous model to study the interaction of energy and the rest of the economy. See their paper in *Modeling Energy Economy Interactions: Five Approaches*, ed. Charles Hitch (Resources for the Future, 1978).

$F(K, L, R)$ were a Cobb–Douglas (or, more generally, of the form $f(K, L) R^s$), then s would be the usual Cobb–Douglas exponent for the resource input. If we follow the usual practice of using the share of the national income imputed to natural resources (pR/Q) as a first approximation to s, then it is a small number. As will be seen later, it would have to be put in the range from 0.02 to 0.10 for most developed countries. In that case, the formula says: for every one percent increase in the cost of natural resources, the net output of the economy is reduced by something between five-hundredths and eleven-hundredths of one percent. That certainly suggests that modern industrial economies are relatively invulnerable to at least a gentle rise in the cost of resources. The reason is that resources are not an "important" input to aggregate production; s is clearly one measure of importance, and, by that measure, resources are small potatoes compared to labor, or even to capital.

The simple answer given by the model is not really as simple as it looks. In the first place, the formula is exact only for small percentage changes in p; for larger changes the formula tends to be overpessimistic. That flaw can easily be corrected either by making exact calculations (which is straightforward but tedious) or by considering a steady reduction in resource availability in a growth process (which I shall do later in this paper).

The second problem with the simple formula is that the key number s is not a constant, except in the Cobb–Douglas case. Thus if the economy becomes steadily more resource-poor, the current value of s might change; and, for that matter, its first-approximation empirical counterpart, the imputed share of natural resources in national product, might well change too. Even the simple formula might be improved. In one special case, the formula can indeed be decomposed into more "fundamental" elements.

In technical terms, the special case is that in which (a) R is "separable" from K and L in the production function F and (b) there is a constant elasticity of substitution between R and the appropriate composite index of K and L. It is not an intolerably special special case, as these things go. What is more important, however, is that there is no particular reason to believe that this special case has any intrinsic bias with respect to the question being analyzed.

Let the production function for gross output have this special form:

$$F(K, L, R) = \left[aR^{(\sigma-1)/\sigma} + bC^{(\sigma-1)/\sigma}\right]^{\sigma/(\sigma-1)}$$

where $C = f(K, L)$ is a notation for the composite index of labor and capital inputs. Now σ is the elasticity of substitution between R and C. It is a constant that measures the ease with which R and C can be substituted for one another. At one limit, $\sigma = 0$, R and C can be used efficiently only in fixed proportions. At the other limit, $\sigma = \infty$, a unit of R can be substituted for a unit of C *ad lib* with no loss of output no matter how far substitution has already been carried. The constants a and b (or rather their ratio a/b) are intrinsic measures of the relative "importance" of the inputs R and C. That is more or less intuitive from the way they enter as weights in the production function. More precisely, it can be calculated that for an economy that responds to resource prices efficiently, the "share" of resources in gross output, the important number $s = a^\sigma p^{1-\sigma}$. The point of this additional step is that the value of s is now reduced to given s—the constants a and σ characterizing the technology, and the parametric price ratio p whose variation is to be studied. Notice that in the exceptional case that $\sigma = 1$, which returns us to the Cobb–Douglas production function, s is independent of p and indeed $s = a$. Thus the interpretation of a as an "importance"-parameter is reinforced.

Now we can push the analysis further. It remains true, as before, that the elasticity of net output with respect to p is $-s/(1-s)$. Moreover

(1) For given σ and p, s is bigger or smaller as a is bigger or smaller. Thus a rise in p is indeed more damaging to the economy the more important resources are as an imput.

(2) For given a and p, it can be shown that s is bigger or smaller as σ is smaller or larger. Thus a rise in p is indeed more damaging to the economy the harder it is to substitute labor and capital for nonrenewable resources in aggregate production.

(3) For given a and σ, s gets bigger as p gets bigger if $\sigma < 1$. But if $\sigma > 1$, resources get less important as p rises.

This is an important conclusion. If the elasticity of substitution exceeds one, that is, if resources and other inputs are easily substituted for one another, the importance of resource inputs diminishes as the cost of resources rises. In other words, as resources get scarcer and scarcer, their scarcity becomes less and less capable of damaging the economy's net output. Thus the ultimate damage is self-limiting. In the opposite case, when the elasticity of substitution is less than one, successive percentage cost increases generate successively larger percentage decreases in net output. This qualitative description suggests a sharper dichotomy than is actually the case. Numerical calculation suggests that with the current approximate value for s, the economy's vulnerability to even drastically reduced resource supplies is quite small unless σ is as low as 0.3 and this, as will be seen, seems to be very far from the "facts."

III. The Problem in a Growth Setting

That is about as far as static theory will carry us. It is quite far, because one could simulate the effects of a progressive increase in resource costs by a succession of static computations, each time allowing for whatever change in K and L is expected to occur. One could even—as suggested by Hogan and Manne—allow for the difference between short-run and long-run elasticities of substitution by decomposing the response of the economy into a chain of steps during which σ goes from its smaller short-run value to its larger long-run value. One could even manage to allow for technical progress in the course of such a sequence of steps. But it is easier instead to embed the model developed so far in the conventional growth-theory framework. In this context it is far simpler—though not logically necessary—to deal with gross output instead of net output, and to represent encroaching resource scarcity just by a reduction in actual resource input. The rise in resource costs that motivates the process can be patched into the end result later.

For this purpose I return to the production function used earlier, according to which

$$Q = \left[a(t)\,R^{(\sigma-1)/\sigma} + b(t)\,C^{(\sigma-1)/\sigma}\right]^{\sigma/(\sigma-1)}.$$

Only now I specify that a and b are functions of time, to allow for the possibility of *technological progress*. In the present notation, the ratio $(a/b)^{(\sigma-1)/\sigma}$ measures the relative importance of natural resources as an input. If that ratio remains constant through time—while a and b both increase proportionally, technical progress is said to be (Hicks-) neutral. Generally an increase in a is said to constitute resource-augmenting technological progress, while a rise in b is, in this special case, C-augmenting, where $C = f(K, L)$ is the composite labor-and-capital input. (I should perhaps emphasize that I use this arbitrary composite only to save time and trouble; since the emphasis is on R, there is no particular need to distinguish between K and L, so I group them in the composite C.)

A well-known decomposition tells us that

$$g_Q = s(g_a + g_R) + (1 - s)(g_b + g_C),$$

where a symbol like g_x stands for the proportional rate of growth of x. Thus the rate of growth of gross output is a weighted average of the growth rates of "effective" resource input and of "effective" labor-and-capital input. The weights are the share elasticities discussed earlier. The growth rate of effective resource input is the sum of the growth rate of actual resource input and the rate of resource-augmenting technological progress. Similarly for the effective composite C. The point of this is that if, over a decade, say, the flow of resource input stays constant ($g_R = 0$) but resource-augmenting technological progress goes on at a rate of one percent a year ($g_a = 0.01$/yr) then, so far as gross output is concerned, it is as if actual resource input were rising at one percent annually with unchanged technology. Or, to take another illustration, a fall in the consumption of natural resources might in principle be exactly offset by an equal percentage resource-augmenting improvement in technology. (Caution: resource-augmenting is not the same thing as resource-saving.)

In fact, since we know how to reduce the share-elasticity s to a function of the given parameters a, σ, and p, this formula tells us in principle most of what we want to know. "All" that is needed is information about a, σ and p. Before we turn to see what the data suggest about plausible values for the parameters, it is worthwhile to list the qualitative conclusions that can be read off from the theoretical formula itself.

(1) A reduction in the flow of nonrenewable resources (i.e., a negative value for g_R) can be offset by resource-augmenting technological progress (a positive g_a) or by labor-and-capital-augmenting technological progress (positive g_b) or by growth of labor and/or capital input (positive g_C).

(2) So long as s is small (like 0.05–0.10), a reduction of the flow of resources is not likely to have a major effect on the growth of gross output, and probably less on the growth of net output.

(3) So long as s is small, a small increase in the rate of growth of the nonresource inputs is enough to offset even a large reduction in the growth-rate of resource inputs.

(4) The economy becomes more vulnerable to resource-scarcity if s increases in the course of time. If, as seems likely, the rate of growth of effective labor-and-capital exceeds the rate of growth of resource input, then s will increase if the elasticity of substitution is less than one, and this increase will have dramatic effects on output only if the elasticity of substitution is considerably less than one.

IV. Some Stylized Facts

For a grasp of orders of magnitude, it is pretty clear that the key number is the share-elasticity *s*. Here are Edward Denison's estimates of the share of the national income inputed to "land" in the U.S.

TABLE 1:

Period	Land share of national income in percent
1909–13	8.9
1914–18	8.8
1919–23	7.0
1924–28	6.4
1929–33	6.2
1934–38	5.6
1939–43	4.9
1944–48	4.0
1949–53	3.4
1954–58	3.0
1909–58	5.8
1909–29	7.7
1929–58	4.5

(For 1930–40 and 1942–46, interpolated rather than actual distributions)

Source: E. F. Denison, *The Sources of Economic Growth in the U.S.*, 1962, p. 30.

TABLE 2:

Year	Land share in nonresidential business in percent
1929	5.36
1940	4.03
1941	4.83
1947	4.47
1948	4.95
1949	4.26
1950	4.46
1951	4.57
1952	4.02
1953	3.58
1954	3.44
1955	3.76
1956	3.39
1957	3.17
1958	3.10
1959	3.35
1960	3.28
1961	3.39
1962	3.71
1963	3.85
1964	3.95
1965	4.40
1966	4.50
1967	4.10
1968	4.04

Source: E. F. Denison, *Accounting for U.S. Economic Growth 1929–1969*, 1974, p. 260.

It takes no fine eye to see, first, that the share-elasticity for land has run between .03 and .09 since the turn of the century and, second, that the trend has been downward from 1909 to 1958, with a gentle rise back to .04 from 1958 to 1968.

Unfortunately, Denison's series is not exactly our *s*. In his interpretation "land" is *defined* as a factor of production whose input remains constant over time. Its main component, especially early in the century, is agricultural and industrial land. Now agricultural and industrial land are natural resources, all right, but they are renewable natural resources. As such, they are not what the fuss is about. Before I give some more appropriate figures, it is worth mentioning that Denison's survey of Europe in the period 1950–1962 gives results entirely compatible with his findings for the U.S. On average, the share of "land" in nonresidential national income ranged from 2.9 percent in the U.K. through 4.0 percent in Northwest Europe to 6.6 percent in Italy. Denison's figure for Japan ranges around 4.5 to 5.0 percent. (The precisely corresponding number for the U.S. was 3.0 percent.)

"Land" however is not exactly what we are after. Table 3 gives a better approximation to the relevant share-elasticity *s*. It records the ratio of the value of mineral resource inputs to the gross national product annually since 1900. The mineral resources covered include mineral fuels, metals and non-metallic minerals. No nonrenewable resources of importance are omitted. The appropriate measure of resource input seems to be "primary demand"—domestic mine production plus imports minus exports—measured in current prices. I have used current-price GNP as the denominator, though perhaps national income (i.e., net national product at factor cost, more or less) would be more appropriate.

The message of Table 3 is clear. There are some short-run irregularities. The ratio *s* falls in the depression of the 1930s, probably because the output of durable goods falls, and also during the

TABLE 3:
Ratio of Nonrenewable Resource Input to GNP

1973		.032	1948	.048	1923	.059
1972		.030	1947	.042	1922	.058
1971		.031	1946	.035	1921	.053
1970		.032	1945	.031	1920	.065
1969	.035	.031	1944	.031	1919	.048
1968	.034		1943	.032	1918	.056
1967	.035		1942	.037	1917	.064
1966	.034		1941	.044	1916	.059
1965	.037		1940	.040	1915	.048
1964	.038		1939	.040	1914	.044
1963	.039		1938	.038	1913	.049
1962	.039		1937	.044	1912	.046
1961	.040		1936	.043	1911	.044
1960	.041		1935	.040	1910	.045
1959	.042		1934	.041	1909	.043
1958	.043		1933	.036	1908	.046
1957	.045		1932	.034	1907	.051
1956	.045		1931	.034	1906	.050
1955	.045		1930	.042	1905	.049
1954	.044		1929	.046	1904	.047
1953	.044		1928	.045	1903	.051
1952	.042		1927	.047	1902	.046
1951	.044		1926	.054	1901	.044
1950	.047		1925	.051	1900	.045
1949	.045		1924	.055		

Source: various issues of *Minerals Yearbook.*

second world war, perhaps because price controls on resource products were more rigid than elsewhere. There are some very brief fluctuations which I am inclined to attribute to short-run supply inelasticity. But generally speaking s averaged about 0.05 in the first thirty years of the century, and rarely fell below 0.045. Since the end of the second war, the value of s has trended slowly downward to about 0.035 or less in the early 1970s. (It will be interesting to bring the table up to date, but I have not yet seen the data for 1974–1976.)

The small size of s is confirmed, apparently beyond doubt. Moreover, if there is anything to the folklore, the U.S. economy is more resource-intensive than most modern industrial economies. There seems to be nothing in this history to contradict the qualitative theoretical conclusions reached earlier; according to the special definition appropriate here, nonrenewable resources have not been an "important" input, and therefore small shifts in the cost gradient would have very small effects on aggregate production. The downward trend in s seems less well established; if it is there, it may date from the end of the first war, or perhaps only from the end of the second.

So much for s. The other important parameter is the elasticity of substitution σ. It is even more of an abstraction than s, and information about σ is even more indirect. The most extensive attempt I know is by Nordhaus and Tobin.* Their procedure is to insert many different specifications of a production function into a simulated growth process (with realistic population growth and saving patterns) and to see which specifications give results that look the most like the stylized facts of U.S. economic growth. Their "best" simulations all had the property that the elasticity of substitution between resources (Denison's "land") and the capital-labor composite was bigger than one. Most of them had a positive value for g_a, usually somewhat smaller than g_b. Always the growth rate of "effective resources" was smaller than that of "effective labor-and-capital." Thus the ratio of effective C to effective R was growing over time. With σ greater than one, it follows that s was falling over time. More accurately, Tobin–Nordhaus conclude that

* "Is Growth Obsolete?", in *National Bureau of Economics Research Fiftieth Anniversary Series*, Vol. V, 1972.

$\sigma > 1$ *because* they give high marks to simulations that show s to be falling over time. They are impelled to do this because they take their cue from Denison's estimates, shown in Tables 1 and 2. The more appropriate data recorded in Table 3 are less forceful on this point, though they do rather confirm it.

Since the end of the second war, real nonrenewable resource input in the U.S. economy has grown at about 3.5 percent annually. In "natural" units, then, resource input R has lately been growing faster than the composite labor-and-capital input C. If Tobin and Nordhaus are right that g_a is less than g_b, then it could still be that the growth rate of "effective resources" has been smaller than that of "effective labor-and-capital" but I am not so sure about that. Therefore I have little confidence in the conclusion that the elasticity of substitution between R and C exceeds one. On the other hand, there is nothing in the figures to suggest that the elasticity of substitution is perceptibly smaller than one,* as it would have to be to create substantial vulnerability in the foreseeable future.

This is all extraordinarily tenuous, but the only possible inference from the work of Denison, Nordhaus–Tobin and Hogan–Manne is that past economic growth would not have gained very much from cheaper or more abundant access to nonrenewable resources, nor lost very much from the opposite. Political events aside, the evidence is that the future will be rather like the past. I do not regard this as a very strong conclusion; but it is safe to say that the opposite conclusion has considerably less evidence or none at all going for it.

* There is some recent econometric work by Ernst Berndt that seems to suggest rather lower elasticities of substitution, but it is not really comparable to the Nordhaus–Tobin exercise. Berndt's estimates are for manufacturing, not for GNP as a whole; they apply to gross output, not value added; and they cover the postwar period, whereas Nordhaus–Tobin try to reproduce longer-run trend behavior. See Berndt and Khaled, "Energy Prices, Economies of Scale and Biased Productivity Gains in U.S. Manufacturing, 1947–1971". University of British Columbia Mimeo, 1977.